Appendix on Mathematical Methods

Contents

This appendix discusses the main mathematical methods used in the text. We consider differential equations, static optimization, dynamic optimization, some results in matrix theory, and a few results from calculus.

A.1 Differential Equations

A.1.1 Introduction

A differential equation is an equation that involves derivatives of variables. If there is only one independent variable, it is called an *ordinary differential equation* (ODE). The *order* of the ODE is that of the highest derivative; that is, if the highest derivative is an ODE of order n, it is an nth-order ODE. When the functional form of the equation is linear, it is a *linear ODE*. Most of the differential equations that we encounter in the book involve derivatives of functions with respect to *time*.

An example of a differential equation is

$$a_1 \cdot \dot{y}(t) + a_2 \cdot y(t) + x(t) = 0 \tag{A.1}$$

where the dot on top of $y(t)$ represents the derivative of $y(t)$ with respect to time, $\dot{y}(t) \equiv dy(t)/dt$, a_1 and a_2 are constants, and $x(t)$ is a known function of time. The function $x(t)$ is sometimes called the *forcing function*. Equation (A.1) is a first-order linear ODE with constant coefficients. If $x(t) = a_3$, a constant, the equation is called *autonomous*. (An equation is autonomous when it depends on time only through the variable $y[t]$.) If $x(t) = 0$, the equation is called *homogeneous*.

A second-order, linear ODE with constant coefficients takes the form,

$$a_1 \cdot \ddot{y}(t) + a_2 \cdot \dot{y}(t) + a_3 \cdot y(t) + x(t) = 0 \tag{A.2}$$

where a_1, a_2, and a_3 are constants and $\ddot{y}(t) \equiv d^2y(t)/dt^2$. The equation

$$a_1 \cdot \dot{y}(t) + a_2(t) \cdot y(t) + x(t) = 0 \tag{A.3}$$

where $a_2(t)$ is a known function of time, is a first-order, linear ODE with *variable coefficients*. The equation

$$\log[\dot{y}(t)] + 1/y(t) = 0 \tag{A.4}$$

is a *nonlinear first-order ODE*.

The goal when solving a differential equation is to find the behavior of $y(t)$. The first solution method that we use is *graphical*, a technique that can be used for nonlinear, as well as linear, differential equations. The disadvantage is that it can be used only for autonomous equations. The second method is *analytical*. In some circumstances, we will be able to find an exact formula for $y(t)$, even when the equation is not autonomous. The drawback of the analytical approach is that it can be used only with a limited set of functional forms. One of them, however, is the linear function in equation (A.1). When we encounter nonlinear differential equations, we will often approximate the solution by linearizing the equation by means of a Taylor-series expansion. (See section A.6.2.)

A third method for solving differential equations relies on numerical analysis. Most modern mathematical computer packages contain subroutines that solve differential equations numerically. Matlab, for example, has the subroutines ODE23 and ODE45, and Mathematica has the command NDSOLVE.

A.1.2 First-Order Ordinary Differential Equations

Graphical Solutions

CONSTRUCTING THE DIAGRAM Consider an autonomous ordinary differential equation of the form,

$$\dot{y}(t) = f[y(t)] \tag{A.5}$$

where $f(\cdot)$ is a known function. Equation (A.5) is autonomous because the function $f(\cdot)$ does not depend on time independently of y. The function $f(\cdot)$ may or may not be linear.

To solve equation (A.5) graphically, we plot $f(\cdot)$ as a function of y in figure A.1. The horizontal axis shows the value of y, and the vertical axis has $f(\cdot)$ and $\dot{y}$. Positive values of $f(\cdot)$ correspond to positive values of $\dot{y}$, in accordance with equation (A.5). Since $\dot{y}$ is the derivative of y with respect to time, positive values of $\dot{y}$ correspond to increasing values of y. To reflect this relation, we draw arrows pointing east (increasing y) when $f(\cdot)$ lies above

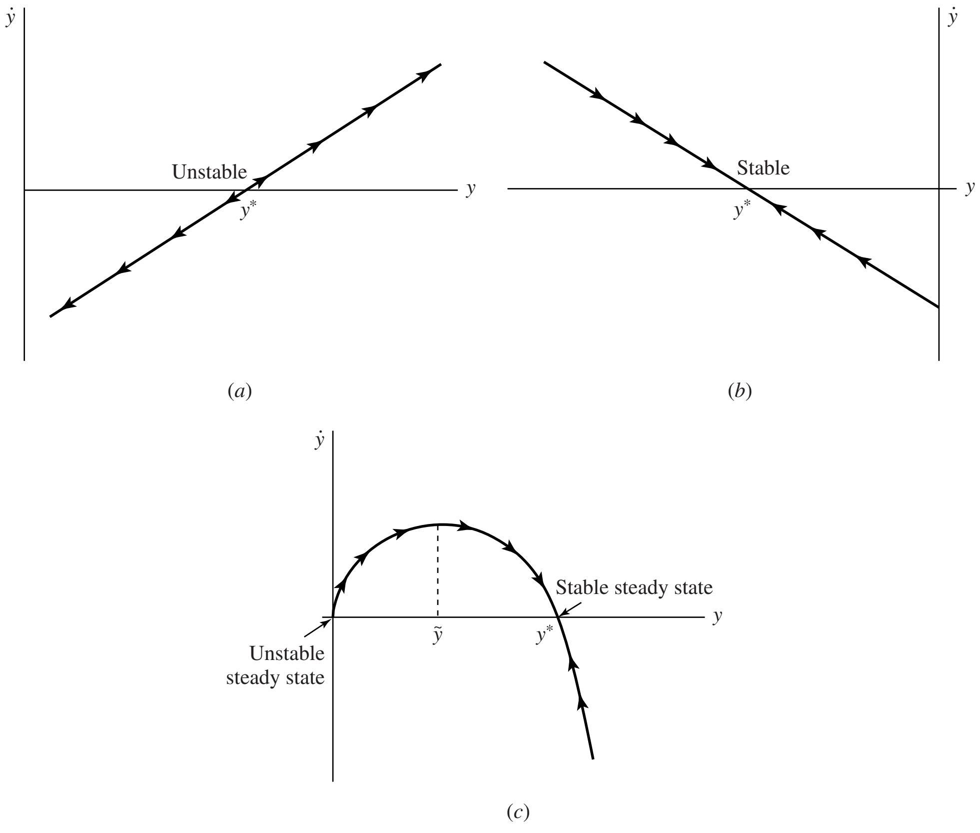


Figure A.1
(*a*) **Linear ODE.** If the coefficient a in equation (A.6) is positive, then the differential equation for y is unstable. (*b*) **Linear ODE.** If the coefficient a in equation (A.6) is negative, then the differential equation for y is stable. (*c*) **Nonlinear ODE.** In equation (A.7), the slope of $f(\bullet)$ with respect to y is initially positive and is subsequently negative. The steady state at 0 is unstable, whereas that at y^* is stable.

the horizontal axis and pointing west (decreasing y) when $f(\cdot)$ lies below the horizontal axis. The arrows reveal the direction in which y moves over time and therefore provide a qualitative solution to the differential equation.

Sometimes the differential equation is expressed in terms of the difference of two functions, for example,

$$\dot{y}(t) = f[y(t)] - g[y(t)]$$

Instead of graphing $f(\cdot) - g(\cdot)$, we can graph $f(\cdot)$ and $g(\cdot)$ separately. The rate of change of $y(t)$, $\dot{y}(t)$, is given in this case by the vertical distance between $f(\cdot)$ and $g(\cdot)$. For values of y where $f(\cdot)$ lies above $g(\cdot)$, $\dot{y}(t)$ is positive and therefore $y(t)$ is increasing over time. The opposite is true when $f(\cdot)$ lies below $g(\cdot)$. The steady state is given by the point(s) at which the curves $f(\cdot)$ and $g(\cdot)$ cross.

As an example, consider a linear form for $f(\cdot)$:

$$\dot{y}(t) = f[y(t)] = a \cdot y(t) - x \tag{A.6}$$

where a and x are constants, with $a > 0$. The graph of $f(\cdot)$ is a straight line with positive slope. This line, depicted in figure A.1a, intercepts the vertical axis at $\dot{y} = -x$ and crosses the horizontal axis at $y^* = x/a$. For values of y above y^*, the function lies above the horizontal axis. Thus, $\dot{y}$ is positive and y is increasing. Hence, to the right of y^*, we draw arrows pointing northeast (see figure A.1a). The opposite conditions apply to the left of y^*, and we draw arrows pointing southwest.

If the initial value, $y(0)$, equals y^*, equation (A.6) implies that $\dot{y}$ equals 0, so that y does not change over time. It follows that $y(t)$ remains forever at y^*. The value y^* is called the *steady-state* value of y.

If $y(0) > y^*$, then $\dot{y} > 0$, so that y grows over time. Conversely, if $y(0) < y^*$, then $\dot{y} < 0$, so that y decreases over time. The qualitative dynamics of $y(t)$ are fully determined in figure A.1a: once the initial value, $y(0)$, is specified, the arrows show how y moves as time evolves. An interesting point is that unless $y(0) = y^*$, the dynamics of the equation when $a > 0$ move y away from the steady state. This behavior applies for initial values below and above y^*. In this case, we say that the differential equation is *unstable*.

Imagine now that $a < 0$. The graph of $f(\cdot)$ is then a downward-sloping straight line, depicted in figure A.1b, which intercepts the vertical axis at $\dot{y} = -x$ and the horizontal axis at $y^* = -x/a$. To the left of y^*, $\dot{y}$ is positive, so that y increases over time. Correspondingly, the arrows in the figure point southeast. The opposite relation applies to the right of y^*. Note that, regardless of the initial value, $y(0)$, the dynamics of the equation brings $y(t)$ back to the steady state, y^*. In this case, we say that equation (A.6) is *stable*.

This graphical approach can be used to analyze the dynamics of more complicated nonlinear functions. Consider, for example, the differential equation

$$\dot{y}(t) = f[y(t)] = s \cdot [y(t)]^{\alpha} - \delta \cdot y(t) \tag{A.7}$$

where s, δ, and α are positive constants and $\alpha < 1$. Chapter 1 shows that the fundamental equation of the Solow–Swan growth model takes the form of equation (A.7), where $y(t)$ is the capital stock. Under this interpretation, equation (A.7) says that the net increase in the capital stock equals the difference between total saving and total depreciation. Total saving is assumed to be the constant fraction, s, of output, y^{α}, and total depreciation is proportional to the existing capital stock.

Since only nonnegative values of the capital stock are economically meaningful, we look only at the first quadrant in figure A.1c. For low values of y, the function $f(\cdot)$ is upward sloping. It reaches a maximum when $s\alpha\tilde{y}^{\alpha-1} = \delta$, and it becomes downward sloping for higher values of y. The function $f(\cdot)$ crosses the horizontal axis at two points, $y = 0$ and $y = y^* = (\delta/s)^{1/(\alpha-1)}$.

To the right of y^*, $\dot{y}$ is negative, so that y is falling. Hence, we draw arrows pointing west. To the left of y^*, $\dot{y}$ is positive, so that y is rising, and we draw arrows pointing east. It follows that the equation has two steady states. The first one is y^* and is stable in that, for any positive initial value, $y(0)$, the dynamics of the equation moves $y(t)$ toward y^*. The second steady state, 0, is unstable: if $y(0) > 0$, the dynamics moves $y(t)$ away from 0.

STABILITY The preceding discussion suggests that if $f(\cdot)$ slopes upward at the steady-state value, y^*, the steady state is unstable. That is, if $y(0) \neq y^*$, $y(t)$ moves away from y^*. The reason is simple: if $f(\cdot)$ is upward sloping when $f(y^*) = 0$, then, for $y > y^*$, $f(y) > 0$. Hence, $\dot{y} > 0$ and y increases over time. On the other hand, for $y < y^*$, $f(y) < 0$, $\dot{y} < 0$, and y decreases over time. The conclusion is that y increases when it is already too large and falls when it is already too small, an indication of instability.

Conversely, if $f(\cdot)$ is downward sloping at the steady-state value, y^*, the equation is stable. In this case, if $y(0) \neq y^*$, $y(t)$ approaches y^* over time.

To summarize, if we are interested in the stability of the differential equation in the neighborhood of a steady state, all we have to do is compute the derivative of $f(\cdot)$ and evaluate it at the steady-state value, y^*:

$$\begin{aligned} &\text{If } \partial\dot{y}/\partial y|_{y^*} > 0,\ y \text{ is locally unstable} \\ &\text{If } \partial\dot{y}/\partial y|_{y^*} < 0,\ y \text{ is locally stable} \end{aligned} \tag{A.8}$$

Although nonlinear differential equations may have more than one steady state, the local stability properties of each of these steady states will still be determined by the condition in equation (A.8).

Analytical Solutions The solution to some equations is almost immediate because the equation can be integrated. For instance, the solution to $\dot{y}(t) = a$ is obviously $y(t) = b + at$, where b is an arbitrary constant.

Equations that involve polynomial functions of time are equally easy to solve, for example,

$$\dot{y}(t) = a_0 + a_1 t + a_2 \cdot t^2 + \cdots + a_n \cdot t^n$$

has the solution

$$y(t) = b + a_0 t + a_1 \cdot (t^2/2) + \cdots + a_n \cdot [t^{n+1}/(n+1)]$$

In general, the functional forms that we work with will not be this simple. We now derive the general solution for linear, first-order ODEs.

LINEAR, FIRST-ORDER DIFFERENTIAL EQUATIONS WITH CONSTANT COEFFICIENTS The general form of the linear, first-order ODE with constant coefficients is

$$\dot{y}(t) + a \cdot y(t) + x(t) = 0 \tag{A.9}$$

where a is a constant and $x(t)$ is a known function of time. The easiest way to solve this equation is to carry out the following steps.

First, put all the terms involving y and its derivatives on one side of the equation and the rest on the other side:

$$\dot{y}(t) + a \cdot y(t) = -x(t)$$

Second, multiply both sides of the equation by e^{at} and integrate:

$$\int e^{at} \cdot [\dot{y}(t) + a \cdot y(t)] \cdot dt = -\int e^{at} \cdot x(t) \cdot dt \tag{A.10}$$

The term e^{at} is called the *integrating factor*. The reason for multiplying by the integrating factor is that the term inside the left-hand side integral becomes the derivative of $e^{at} \cdot y(t)$ with respect to time:

$$e^{at} \cdot [\dot{y}(t) + a \cdot y(t)] = (d/dt)[e^{at} \cdot y(t) + b_0]$$

where b_0 is an arbitrary constant. Note that the integral on the left-hand side of equation (A.10) is the integral of the derivative of some function and therefore equals the function itself (see section A.5.5). Hence, the term on the left-hand side of equation (A.10) equals $e^{at} \cdot y(t) + b_0$.

Third, compute the integral on the right-hand side of equation (A.10), making sure to add another constant term b_1. Note that this integral is a function of t. Call the result $\mathrm{INT}(t) + b_1$. Since $x(t)$ is a known function of time, $\mathrm{INT}(t)$ is also a known function of time.

Fourth, multiply both sides by e^{-at} to get $y(t)$:

$$y(t) = -e^{-at} \cdot \text{INT}(t) + be^{-at} \tag{A.11}$$

where $b = b_1 - b_0$ is an arbitrary constant. Equation (A.11) is the general solution to the ODE in equation (A.9).

Consider, as an example, the differential equation

$$\dot{y}(t) - y(t) - 1 = 0 \tag{A.12}$$

In this example, the forcing function $x(t)$ is a constant, -1. To solve this equation, we follow the steps outlined previously. First, put all the terms involving $y(t)$ and its derivatives on the left-hand side of the equation and all the other terms on the right-hand side. Then multiply both sides by e^{-t} and integrate:

$$\int e^{-t}[\dot{y}(t) - y(t)] \cdot dt = \int e^{-t}\, dt \tag{A.13}$$

The term inside the integral on the left-hand side is the derivative of $e^{-t} \cdot y(t) + b_0$ with respect to time. Hence, the integral on the left-hand side equals $e^{-t} \cdot y(t) + b_0$. The right-hand side equals $-e^{-t} + b_1$. Hence, the solution to equation (A.12) is

$$y(t) = -1 + be^{t} \tag{A.14}$$

where $b = b_1 - b_0$ is an arbitrary constant. We can verify that equation (A.14) satisfies equation (A.12) by taking derivatives with respect to time to get $\dot{y}(t) = be^{t} = y(t) + 1$.

The result in equation (A.11) is the *general solution* to equation (A.9); to get a *particular* or *exact solution,* we have to specify the arbitrary constant of integration, b. To pin down which of the infinitely many possible paths applies, we need to know a value of $y(t)$ for at least one point in time. This *boundary condition* will determine the unique solution to the differential equation.

Figure A.2 shows an array of solutions to the ODE in the example of equation (A.12). To choose among them, imagine that we know that $y(t) = 0$ when $t = 0$. This type of boundary condition is called an *initial condition* because it pins down the path by specifying the value of $y(t)$ at the initial date. In our example, we can substitute $t = 0$ and $y(0) = 0$ in equation (A.14) to find that $y(0) = -1 + be^{0} = 0$, which implies $b = 1$. We can therefore plug $b = 1$ into equation (A.14) to get the particular solution,

$$y(t) = -1 + e^{t} \tag{A.15}$$

This equation, which determines a unique value of y at each point in time, corresponds to the time path labeled A in figure A.2.

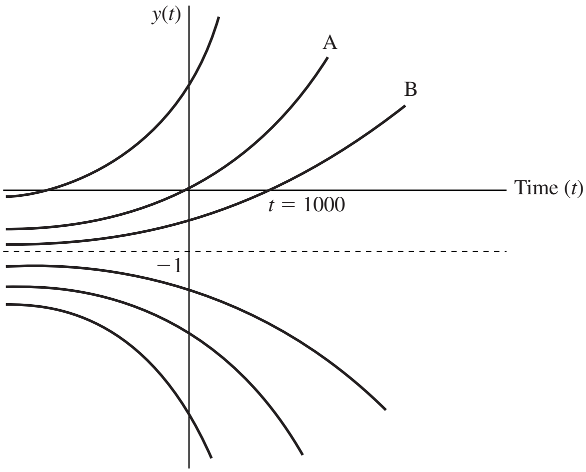


Figure A.2
Solutions to a differential equation. The figure shows an array of solutions to the differential equation (A.12).

Instead of knowing the initial value of the function, we may know the value at some terminal date; that is, we could have a *terminal condition*.[1] As an example, suppose that the terminal date is $t_1 = 1000$, and the value of $y(t)$ at that time is 0. Thus, $y(1000) = -1 + b \cdot e^{1000} = 0$. The solution, $b = e^{-1000}$, implies

$$y(t) = -1 + (e^{-1000}) \cdot e^t \tag{A.16}$$

This result corresponds to path B in figure A.2.

LINEAR, FIRST-ORDER DIFFERENTIAL EQUATIONS WITH VARIABLE COEFFICIENTS Consider now the differential equation

$$\dot{y}(t) + a(t) \cdot y(t) + x(t) = 0 \tag{A.17}$$

where $a(t)$ is a known function of time but is no longer a constant. We can follow the same steps as before. The difference is that the integrating factor is now $e^{\int_0^t a(\tau)d\tau}$, so that the left-hand side becomes the derivative of $y(t) \cdot e^{\int_0^t a(\tau)d\tau}$.[2] Again, when we integrate the derivative of a function, we get back the original function. Using this information, we

1. When we deal with growth models with infinite horizons, we may know the limiting value of a variable as time tends to infinity. This information will provide us with a terminal condition.

2. The lower limit of integration can be an arbitrary constant. Leibniz's rule for differentiation of definite integrals says that $d[\int_0^t f(\tau)\, d\tau]/dt = f(t)$. Note that we are taking the derivative with respect to the upper limit of integration. See section A.6.6.

find that the solution to the ODE is

$$y(t) = -e^{-\int_0^t a(\tau)d\tau} \cdot \int e^{\int_0^t a(\tau)d\tau} \cdot x(t) \cdot dt + b \cdot e^{-\int_0^t a(\tau)d\tau} \tag{A.18}$$

where b is an arbitrary constant of integration. To find the particular or exact solution, we again have to make use of a boundary condition.

A.1.3 Systems of Linear Ordinary Differential Equations

We now study a system of linear, first-order ODEs of the form

$$\dot{y}_1(t) = a_{11}y_1(t) + \cdots + a_{1n}y_n(t) + x_1(t)$$

$$\cdots$$

$$\dot{y}_n(t) = a_{n1}y_1(t) + \cdots + a_{nn}y_n(t) + x_n(t)$$

In matrix notation, the system is

$$\dot{y}(t) = A \cdot y(t) + x(t) \tag{A.19}$$

where $y(t)$ is a column vector of n functions of time, $\begin{bmatrix} y_1(t) \\ \cdots \\ y_n(t) \end{bmatrix}$, $\dot{y}(t)$ is the column vector of the n corresponding derivatives, A is an $n \times n$ square matrix of constant coefficients, and $x(t)$ is a vector of n functions.

We consider three procedures for solving this system of differential equations. The first one is a graphical device called a *phase diagram,* similar to the one that we used for a single differential equation. The advantage of a phase diagram is that it is simple and provides a qualitative solution. Furthermore, this technique works for nonlinear, as well as linear, systems. The drawbacks of phase diagrams are that they work only for 2×2 systems and only for autonomous equations with steady states.

The second procedure is *analytical*. The advantages of the analytical approach are that it gives quantitative answers and can be used in larger systems. The disadvantage is that it works, in general, only for linear equations. Later in this section, however, we use linear approximations to nonlinear systems.

The third procedure is *numerical*. Later in this section, we describe the time-elimination method for solving nonlinear systems numerically.

Phase Diagrams

DIAGONAL SYSTEMS Start with the simple case in which A is a 2×2 *diagonal* matrix and the equations are homogeneous; that is, the components of the vector $x(t)$ are 0. The system

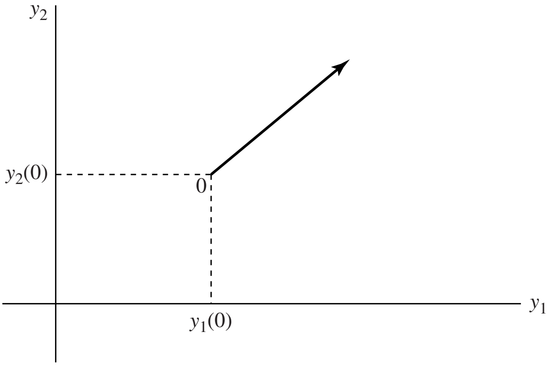


Figure A.3
Directions of motion. The figure shows the directions of motion for y_1 and y_2 in the diagonal system given in equation (A.20).

can then be rewritten as

$$\dot{y}_1(t) = a_{11} \cdot y_1(t)$$
$$\dot{y}_2(t) = a_{22} \cdot y_2(t) \tag{A.20}$$

where a_{11} and a_{22} are real numbers.

A phase diagram is a graphical tool, similar to the one used in the previous section, which allows us to visualize the dynamics of the system. In figure A.3, y_1 is on the horizontal axis, and y_2 is on the vertical axis. Each point in the space represents the position of the system (y_1, y_2) at a given moment in time. Imagine that, at time 0, we are at the point labeled "0" in the figure; that is, y_1 equals $y_1(0)$ and y_2 equals $y_2(0)$. If we want to see what the position of the economy will be at the "next instant," we could have a third dimension to represent time. More conveniently, we can represent the dynamics with arrows that point in the direction of motion, just as in section A.1.2. For instance, an arrow that points northeast at point "0" signifies that the variables y_1 and y_2 are each growing over time. If the arrow points north, y_2 grows and y_1 is stationary, and so on.

The object of a phase diagram is to translate the dynamics implied by the two differential equations into a system of arrows that describe the qualitative behavior of the economy over time. As a simple example, consider the diagonal system that we studied before. The dynamics depend on the signs of the two diagonal elements of A. We now consider three cases.

Case 1, $a_{11} > 0$ *and* $a_{22} > 0$ To construct the phase diagram, follow the following steps:

1. Start in figure A.4a by plotting the locus of points for which $\dot{y}_1$ equals 0, called the $\dot{y}_1 = 0$ *schedule*. In this case, the locus corresponds to the points for which $y_1(t) = 0$; that is, the vertical axis.

2. Analyze the dynamics of y_1 in each of the two regions generated by the $\dot{y}_1 = 0$ schedule. For positive y_1 (that is, to the right of the $\dot{y}_1 = 0$ schedule), $\dot{y}_1$ is positive because $a_{11} > 0$ and $y_1 > 0$. Hence, the arrows point east. The opposite is true to the left of the vertical axis because in that region, $\dot{y}_1$ is given by the product of a positive number, $a_{11} > 0$, and a negative number, $y_1 < 0$. Therefore, the arrows point west.

3. Repeat the procedure for $\dot{y}_2$. In the present example, the $\dot{y}_2 = 0$ schedule is the horizontal axis shown in figure A.4b. For positive y_2, $\dot{y}_2$ is the product of two positive numbers and is therefore positive. Hence, y_2 is increasing and, correspondingly, the arrows point north. Similarly, the arrows point south for negative y_2.

4. Join the two pictures in figure A.4c. The two schedules divide the space into four regions. (In this simple case, the regions correspond to the four quadrants, a result that is not general.) In the first quadrant, one arrow points east and the other points north. We combine the two into an arrow that points northeast. This construction means that, if the economy is in this region, y_1 and y_2 are increasing. The combined arrows for the second, third, and fourth quadrants point northwest, southwest, and southeast, respectively. Along the vertical axis, the arrows point north for positive y_2 and south for negative y_2. On the horizontal axis, the arrows point east for positive y_1 and west for negative y_1. Finally, at the origin, $\dot{y}_1$ and $\dot{y}_2$ are 0. Hence, if the economy happens to be at the origin, it remains there forever. This point is the *steady state*. It is *unstable* in that if the initial position deviates from the origin by a small amount in any direction, the dynamics of the system (the arrows) take it away from the steady state.

5. Use the boundary conditions to see which one of the many paths depicted in the picture constitutes the exact solution. Imagine, for example, that, at time zero, the value of y_1 is 1 and the value of y_2 is 2. (In this case, the two boundary conditions are initial conditions, but, in other cases that we consider, we may have one initial condition and one terminal condition or two terminal conditions.) The initial conditions imply that the system starts at point "0" in figure A.4c. The subsequent behavior of y_1 and y_2 is given by the path going through "0," as depicted in figure A.4c.

Case 2, $a_{11} < 0$ *and* $a_{22} < 0$ Arguments similar to those of the previous section imply that the $\dot{y}_1 = 0$ schedule is again the vertical axis, and the $\dot{y}_2 = 0$ schedule is again the horizontal axis. We follow the same steps as before to find in figure A.5 that the arrows point southwest in the first quadrant, southeast in the second, northeast in the third, and northwest in the

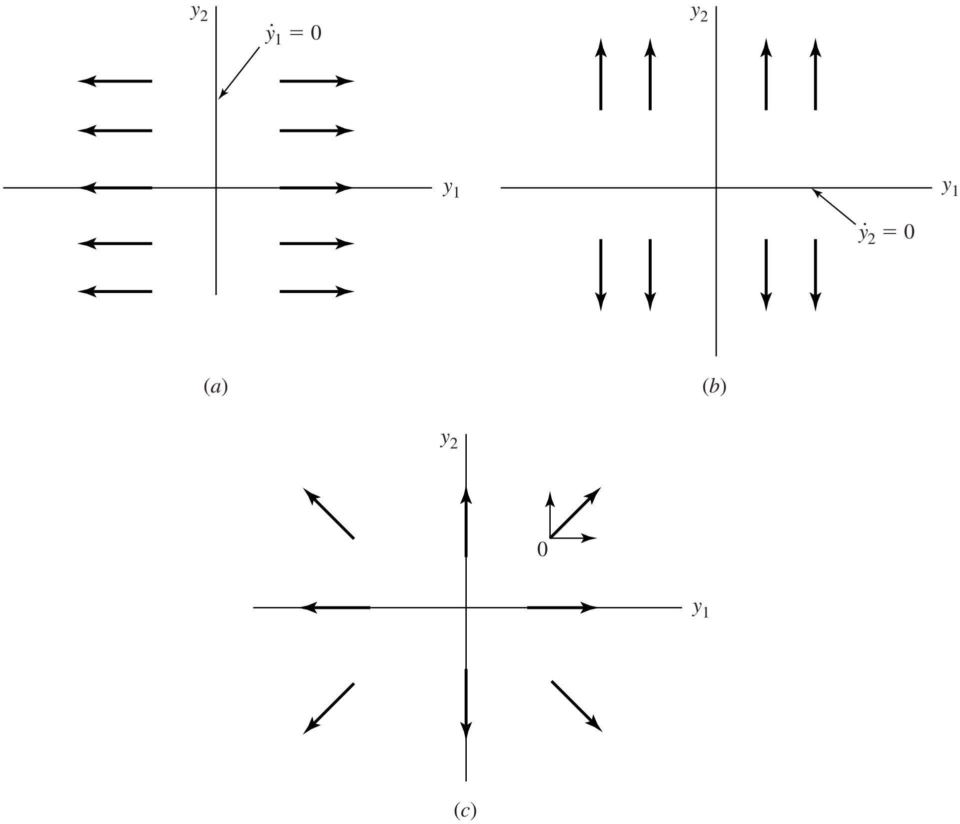


Figure A.4
(*a*) **The $\dot{y}_1 = 0$ locus.** The figure shows the $\dot{y}_1 = 0$ schedule (the vertical axis in this example) for the system in equation (A.20) when $a_{11} > 0$. The arrows show the direction of motion for y_1. (*b*) **The $\dot{y}_2 = 0$ locus.** The figure shows the $\dot{y}_2 = 0$ schedule (the horizontal axis in this example) for the system in equation (A.20) when $a_{22} > 0$. The arrows show the direction of motion for y_2. (*c*) **The phase diagram in an unstable case.** The results from figures A.4(a) and A.4(b) are joined to generate a simple phase diagram. The arrows show the directions of motion for y_1 and y_2 when $a_{11} > 0$ and $a_{22} > 0$. This system is unstable.

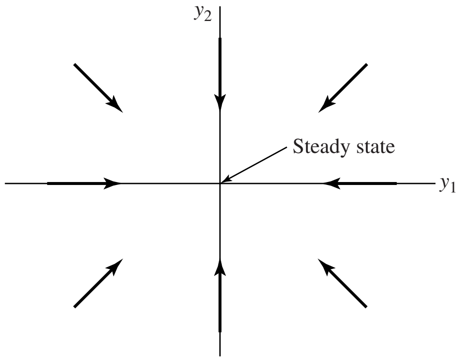


Figure A.5
The phase diagram in a stable case. In this example, $a_{11} < 0$ and $a_{22} < 0$ apply in equation (A.20). This system is stable.

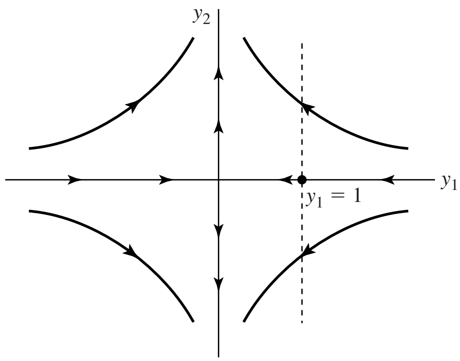


Figure A.6
The phase diagram in a case of saddle-path stability. In this example, $a_{11} < 0$ and $a_{22} > 0$ apply in equation (A.20). This system is saddle-path stable.

fourth. The steady state is the origin and, unlike the previous case, this position is *stable*. For any initial values of y_1 and y_2, the dynamics of the system takes it back to the steady state.

Case 3, $a_{11} < 0$ and $a_{22} > 0$ As in the previous cases, the $\dot{y}_1 = 0$ schedule is the vertical axis, and the $\dot{y}_2 = 0$ schedule is the horizontal axis. The dynamics in this third case, shown in figure A.6, is, however, more complicated than before. The arrows point northwest in the first quadrant, northeast in the second, southeast in the third, and southwest in the fourth. The arrows point toward the origin along the horizontal axis and away from it along the vertical axis. The origin is, again, the steady state.

The new element is that the system is neither stable nor unstable. If the system starts at the steady state, it remains there. If it starts along the horizontal axis, the dynamics of the system takes it back to the steady state. But if the system starts at any point off the horizontal axis, no matter how close to it, the dynamics takes it away from the steady state. The system explodes in the sense that y_2 approaches infinity as t tends to infinity.

This case is called *saddle-path stable*. The reason for this name is the analogy with a marble left on top of a saddle. There is one point on the saddle where, if left there, the marble does not move. This point corresponds to the steady state. There is a trajectory on the saddle with the property that if the marble is left at any point on that trajectory, it rolls toward the steady state. But if the marble is left at any other point, the marble falls to the ground.

Two results about the dynamic paths shown in figure A.6 are worth highlighting. First, none of the paths cross each other. Second, there are only two paths going through the steady state, one is the saddle path that we just mentioned, and the other is the unstable path that corresponds to the vertical axis. These paths are called the *stable arm* and the *unstable arm,* respectively. All two-dimensional systems of ODEs that exhibit saddle-path stability have one stable arm and one unstable arm, each going through the steady state.

Figure A.6 shows the dynamics of the economy for all possible points. The particular path followed depends on two boundary conditions, which have to be specified. As an example, suppose that the initial condition is $y_1(0) = 1$, and the terminal condition is $\lim_{t\to\infty}[y_2(t)] = 0$. The initial condition says that the economy starts anywhere on the vertical line $y_1 = 1$ (see figure A.6). Among all the possible points on this line, only the one on the horizontal axis has the property that y_2 approaches 0 as time goes to infinity. Hence, the terminal condition ensures that the starting point for this economy is $y_2(0) = 0$, right on the stable arm.

By symmetry, the case in which $a_{11} > 0$ and $a_{22} < 0$ also displays saddle-path stability. The only difference is that now the horizontal axis is unstable, whereas the vertical axis is stable.

The key lesson in this section is that if the matrix associated with the system of ODEs is diagonal, its stability properties depend on the signs of the coefficients. If both are positive, the system is unstable. If both are negative, the system is stable. If they have opposite signs, the system is saddle-path stable.

A NONDIAGONAL EXAMPLE When the system of ODEs is nondiagonal, we follow the same steps to construct the phase diagram. As an example, consider the case

$$\begin{aligned} \dot{y}_1(t) &= 0.06 \cdot y_1(t) - y_2(t) + 1.4 \\ \dot{y}_2(t) &= -0.004 \cdot y_1(t) + 0.04 \end{aligned} \tag{A.21}$$

with the boundary conditions $y_1(0) = 1$ and $\lim_{t\to\infty}[e^{-0.06t} \cdot y_1(t)] = 0$.

The $\dot{y}_1 = 0$ locus is the upward-sloping line $y_2 = 1.4 + 0.06 \cdot y_1$. If we start at a point on the $\dot{y}_1 = 0$ schedule and increase y_1 a bit, the right-hand side of the expression for $\dot{y}_1$ in equation (A.21) increases. Hence, $\dot{y}_1$ becomes positive and y_1 is increasing in that region. The arrows in this region therefore point east. A symmetric argument implies that the arrows point west for points to the left of the $\dot{y}_1 = 0$ schedule.

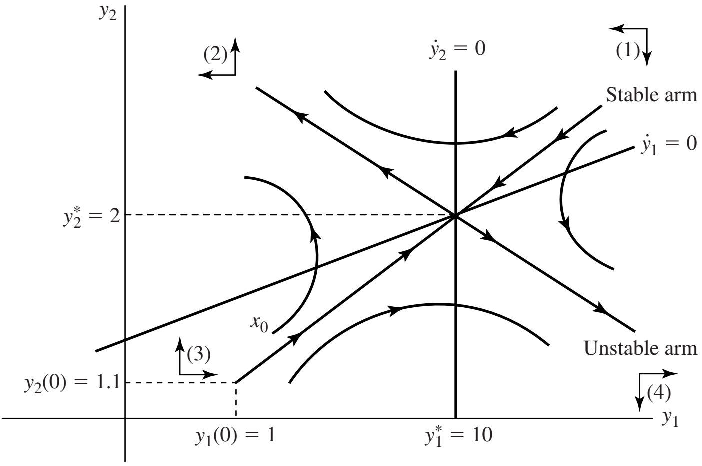
The $\dot{y}_2 = 0$ locus is given by $y_1 = 10$, a vertical line; that is, this locus is independent of y_2. The expression for $\dot{y}_2$ in equation (A.21) implies that if y_1 rises, $\dot{y}_2$ decreases. Hence, to the right of the $\dot{y}_2 = 0$ locus, $\dot{y}_2$ is negative, and the arrows point south. The reverse is true to the left of the locus.

The two loci divide the space into four regions, labeled 1 through 4 in figure A.7a. The steady state is the point at which the two loci cross, a condition that corresponds in this case to $y_1^* = 10$ and $y_2^* = 2$. In region 1, the combined arrows point southwest; in region 2, northwest; in region 3, northeast; and in region 4, southeast.

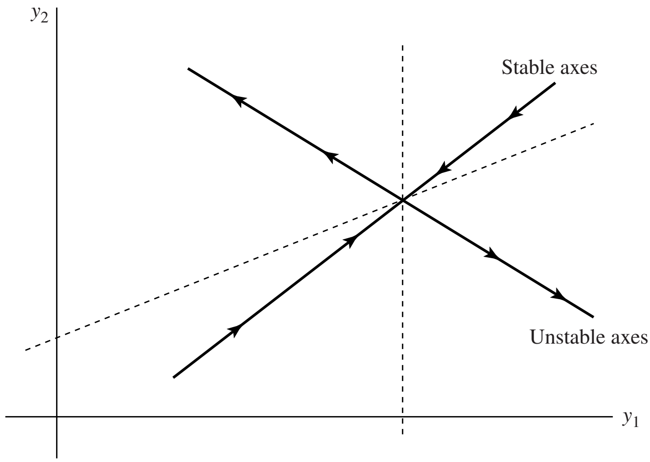
To assess the stability properties of the system, we can ask the following question: From how many of the four regions do the arrows allow the system to move toward the steady state? If the answer is two, the system is saddle-path stable, and the saddle path is located in these two regions.

Figure A.7a shows that the system can move toward the steady state if and only if it starts in regions 1 and 3. Therefore, the system is saddle-path stable. The saddle path, located in regions 1 and 3, goes through the steady state. If the system starts on this path, it converges to the steady state. If it starts slightly above the saddle path in region 3—say, at point x_0 in figure A.7a—it follows the arrows northeast for a while. The path eventually crosses the $\dot{y}_1 = 0$ locus, and the system then moves northwest, away from the steady state. We can also show readily that the system diverges from the steady state if it starts below the stable arm in region 3. In fact, the system diverges from the steady state if it begins at any point that is not on the stable arm.

The exact path along which the system evolves depends on the boundary conditions. This example specifies one initial and one terminal condition. The initial condition says that the system starts somewhere on the vertical line $y_1 = 1$. The terminal condition says that the product of y_1 and a term that goes to 0 at a rate of 0.06 per year goes to 0 as t goes to infinity. If the system ends up in the steady state, y_1 will be constant, so that the product of a constant and a term that approaches zero will be zero. Hence, the terminal condition will be satisfied if y_1 approaches a constant in the long run. If the system does not end up in the steady state, y_1 will increase or decrease at an ever-increasing rate. (The arrows move the economy away from the $\dot{y}_1 = 0$ axis, and y_1 grows in magnitude at an increasing rate.) Since the product of a factor that decreases at rate of 0.06 per year and a factor whose absolute value grows at ever-increasing rates is not 0, the terminal condition requires the system to end up at the steady state. It follows that, because $y_1(0)$ is not at the steady state,



(a)



(b)

Figure A.7
(a) **The phase diagram in a nonlinear example with saddle-path stability.** The figure shows the phase diagram for the system in equation (A.21). This system is saddle-path stable. (b) **The stable arm and the unstable arm.** This figure is generated by erasing the $y_1 = 0$ and $y_2 = 0$ schedules and the normal axes in figure A.7a. We are left with the stable arm and the unstable arm.

the corresponding value $y_2(0)$ must be the one that puts the system on the stable arm, as shown in figure A.7a.

Suppose that we erase the normal axes and the $\dot{y}_1 = 0$ and $\dot{y}_2 = 0$ schedules, as shown in figure A.7b. We are then left with the stable arm (with arrows pointing toward the steady state) and the unstable arm (with arrows pointing away from the steady state). These two lines divide the space into four regions with the corresponding dynamics as represented by the arrows. Note the similarity between figure A.7b and figure A.6. We can, in fact, think of figure A.7b as a distorted version of figure A.6. This perspective will allow us to interpret the analytical solution to these systems.

A NONLINEAR EXAMPLE We conclude this section on phase diagrams with a nonlinear example. Consider the following system:

$$\dot{k}(t) = k(t)^{0.3} - c(t) \tag{A.22}$$

$$\dot{c}(t) = c(t) \cdot [0.3 \cdot k(t)^{-0.7} - 0.06] \tag{A.23}$$

with boundary conditions $k(0) = 1$ and $\lim_{t\to\infty}[e^{-0.06t} \cdot k(t)] = 0$. The main difference between this system and the ones already considered is that the functional forms are now nonlinear. However, to construct a phase diagram for nonlinear systems, we follow exactly the same steps as before.

The $\dot{k} = 0$ locus is given from equation (A.22) by $c = k^{0.3}$. If we put k on the horizontal axis and c on the vertical, this locus is an upward-sloping and concave curve, as shown in figure A.8. Consider a point slightly to the right of the $\dot{k} = 0$ locus; that is, with slightly higher k and the same c. Equation (A.22) implies that the new point has a larger right-hand side; hence, $\dot{k}$ must be positive. Therefore, k rises to the right of the $\dot{k} = 0$ schedule and the arrows point east. A symmetric argument shows that the arrows point west to the left of the $\dot{k} = 0$ schedule.

The $\dot{c} = 0$ schedule is given from equation (A.23) by $k = 10$, a vertical line (see figure A.8). Consider a point to the right of the $\dot{c} = 0$ locus; that is, with the same c and higher k. Equation (A.23) implies $\dot{c} < 0$; hence, the arrows point south. By a similar argument, the arrows to the left of the $\dot{c} = 0$ schedule point north.

We can now combine the dynamics for k and c. The steady state is the point at which the $\dot{k} = 0$ and $\dot{c} = 0$ loci cross, a condition that corresponds to $k^* = 10$ and $c^* = 2$. Figure A.8 shows that the arrows are such that the system approaches the steady state only from regions 1 and 3. We conclude that the system is saddle-path stable. The stable arm in this case is *not* a linear function. It is still true, however, that the stable arm runs between regions 1 and 3 and goes through the steady state. The unstable arm moves between regions 2 and 4.

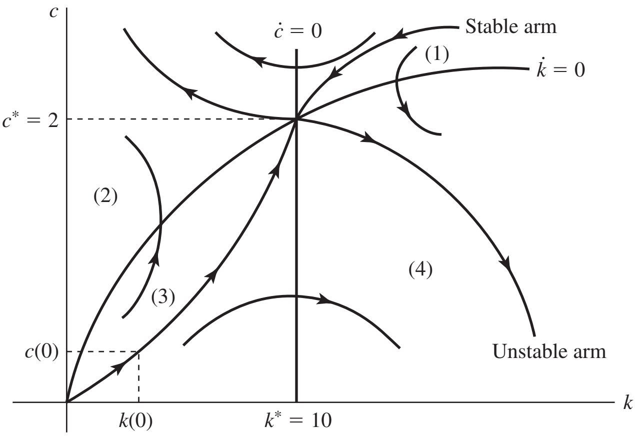


Figure A.8
The phase diagram for a nonlinear model. The figure shows the phase diagram for the system in equations (A.22) and (A.23). This system is saddle-path stable.

We can again use the boundary conditions to select the path that the system will follow. In this example, the boundary conditions ensure that the system begins on the stable arm and therefore approaches its steady state over time.

Analytical Solutions of Linear, Homogeneous Systems We now consider the analytical solution to systems of linear ODEs. We start with the homogeneous case because the solution to the general case is intensive in notation. The $x(t)$ vector in equation (A.19) is then set to 0, so the system becomes

$$\dot{y}(t) = A \cdot y(t) \tag{A.24}$$

where $y(t)$ is an $n \times 1$ column vector of functions of time, $y_i(t)$, A is an $n \times n$ matrix of constant coefficients, and $\dot{y}(t)$ is the vector of time derivatives corresponding to $y(t)$.

Imagine that there is an $n \times n$ matrix V with the property that if we premultiply A by V^{-1} and postmultiply by V, we get a diagonal $n \times n$ matrix:

$$V^{-1}AV = D \tag{A.25}$$

where D is a square matrix in which all the off-diagonal elements are 0. Section A.5 shows that V and D may exist: they are, respectively, the matrix of eigenvectors and the diagonal matrix of eigenvalues associated with A.[3]

3. A sufficient condition for the matrix A to be diagonalizable is for all the eigenvalues to be different. In this case, the eigenvectors are linearly independent, so that $\det(V) \neq 0$ and V^{-1} exists.

We can define the variables $z(t)$ as

$$z(t) = V^{-1} \cdot y(t)$$

Since V^{-1} is a matrix of constants, $\dot{z}(t) = V^{-1} \cdot \dot{y}(t)$. We can therefore rewrite the system from equation (A.24) in terms of the transformed $z(t)$ variables:

$$\dot{z}(t) = V^{-1} \cdot \dot{y}(t) = V^{-1}A \cdot y(t) = V^{-1}AVV^{-1} \cdot y(t) = D \cdot z(t) \tag{A.26}$$

This system consists of n one-dimensional differential equations:

$$\begin{aligned} \dot{z}_1(t) &= \alpha_1 \cdot z_1(t) \\ \dot{z}_2(t) &= \alpha_2 \cdot z_2(t) \\ &\cdots \\ \dot{z}_n(t) &= \alpha_n \cdot z_n(t) \end{aligned} \tag{A.27}$$

We showed in section A.2.2 that the solution for each of these differential equations takes the form $z_i(t) = b_i \cdot e^{\alpha_i t}$, where each b_i is an arbitrary constant of integration that is determined by the boundary conditions (see equation [A.11]). We can express this result in matrix notation as

$$z(t) = Eb \tag{A.28}$$

where E is a diagonal matrix with $e^{\alpha_i t}$ in the ith diagonal term, and b is a column vector of the constants b_i.

We can transform the solution for the z variables back to the y variables by using the relation $y = Vz$. The solution for y is

$$y = VEb$$

or, in nonmatrix notation,

$$y_i(t) = v_{i1}e^{\alpha_1 t} \cdot b_1 + v_{i2}e^{\alpha_2 t} \cdot b_2 + \cdots + v_{in}e^{\alpha_n t} \cdot b_n \tag{A.29}$$

In summary, the general method to solve a system of equations of the form of equation (A.24) is as follows:

1. Find the eigenvalues of the matrix A and call them $\alpha_1, \ldots, \alpha_n$.
2. Find the corresponding eigenvectors and arrange them as columns in a matrix V.
3. The solution takes the form of equation (A.29).
4. Use the boundary conditions to determine the arbitrary constants of integration (b_i).

The Relation Between the Graphical and Analytical Solutions We now relate the graphical and analytical approaches to each other. Remember that when we constructed the phase diagram, we suggested that if we erase the axes and the $\dot{y}_i = 0$ loci and look at the remaining picture in figure A.7b, we get a distorted version of the picture in figure A.6, for which the matrix A was diagonal. We saw also that the analytical solution involved a diagonal matrix of eigenvalues. The similarities in the two approaches are no coincidence: when we diagonalize a matrix we implicitly find a set of axes (or vector basis) on which the linear application represented by A can be expressed as a diagonal matrix (see section A.5). The new axes are the eigenvectors, and the elements in the corresponding diagonal matrix are the eigenvalues.

The graphical solution to the system of equations is basically the same thing. The stable and unstable arms correspond to the two eigenvectors. If we think of these two arms as a new set of axes—that is, if we erase the old axes and the $\dot{y}_i = 0$ schedules—then the old matrix A can be represented by the diagonal eigenvalue matrix. The phase diagram for the nondiagonal case looks accordingly like a distorted version of the diagonal one.

Stability Recall that the stability properties of the diagonal examples depend on the signs of the diagonal elements. Not surprisingly, therefore, the stability properties of the nondiagonal system depend on the signs of its eigenvalues. Several possibilities arise:

1. The two eigenvalues are *real* and *positive*. In this case, the system is unstable.

2. The two eigenvalues are *real* and *negative*. In this case the system is stable.

3. The two eigenvalues are *real* with *opposite signs*. In this case, the system is saddle-path stable. Furthermore, when the system is saddle-path stable, *the stable arm corresponds to the eigenvector associated with the negative eigenvalue*.[4] Similarly, the unstable arm corresponds to the eigenvector associated with the positive eigenvalue. The intuition is again that the axes associated with the diagonal matrix are given by the eigenvectors. As we saw in the examples, when the system is diagonal, the axis associated with the negative component of the diagonal matrix is the stable arm, and the axis associated with the positive component is the unstable arm.

4. The two eigenvalues are *complex* with *negative real parts*. The system converges in this case to the steady state in an oscillating manner (figure A.9a).

5. The two eigenvalues are *complex* with *positive real parts*. The system is unstable and oscillating, as depicted in figure A.9b.

4. Throughout the book we will use interchangeably the terms eigenvector associated with negative eigenvalue and negative eigenvector.

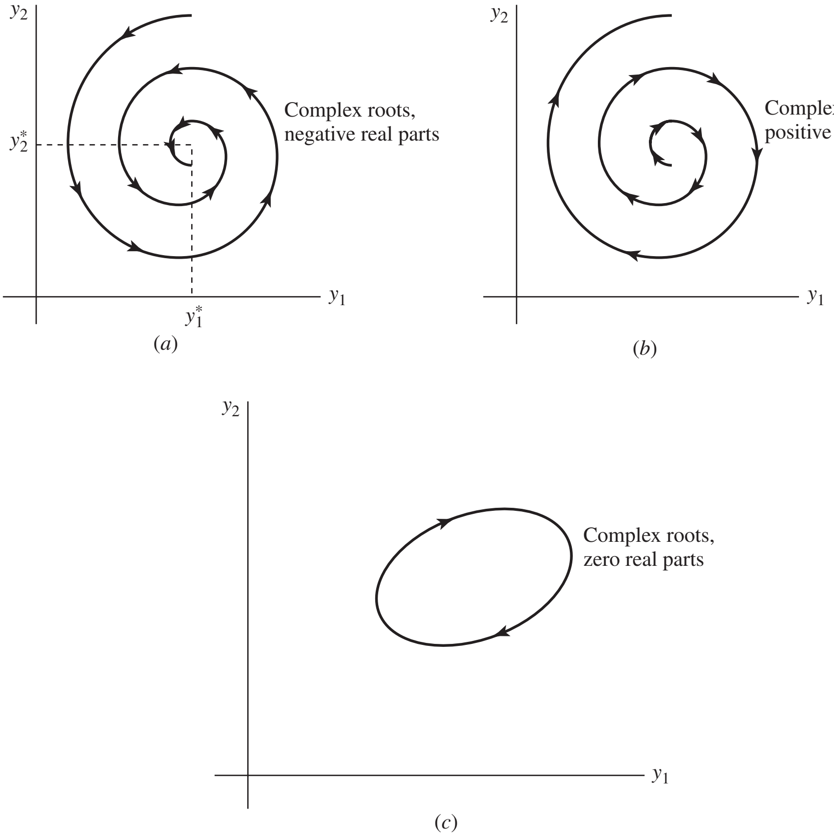


Figure A.9
(*a*) **Stable, oscillating dynamics.** If the two eigenvalues are complex with negative real parts, then the system converges to the steady state in an oscillating manner. (*b*) **Unstable, oscillating dynamics.** If the two eigenvalues are complex with positive real parts, then the system diverges from the steady state in an oscillating manner. (*c*) **Oscillating dynamics.** If the two eigenvalues are complex with 0 real parts, then the trajectories are ellipses around the steady state. This system neither converges nor diverges.

6. The two eigenvalues are *complex* with *zero real parts*. The trajectories are then ellipses around the steady state, as shown in figure A.9c.

7. The two eigenvalues are *equal*. In this case, the matrix of eigenvectors cannot be inverted, and the analytical solution outlined earlier in this section cannot be applied. The solution in this case takes the form

$$y_i(t) = (b_{i1} + b_{i2} \cdot t) \cdot e^{\alpha t}$$

where b_{i1} and b_{i2} are functions of the constants of integration and the coefficients in A, and α is the unique eigenvalue. The solution is stable if $\alpha < 0$ and unstable if $\alpha > 0$.

We should mention that, in nonlinear systems, there is one more type of equilibrium called a *limit cycle*. A stable limit cycle is one toward which trajectories converge, and an unstable limit cycle is one from which trajectories diverge.

The stability properties of systems with higher dimensions are similar. If all eigenvalues are positive, the system is unstable. If all the eigenvalues are negative, the system is stable. If the eigenvalues have different signs, the system is saddle-path stable. Since, as argued before, the stable arm corresponds to the eigenvector(s) associated with the negative eigenvalue(s), the dimension of the stable arm is the number of negative eigenvalues. For instance, in a 3×3 system with one negative eigenvalue, the stable arm is a line going through the steady state and corresponding to the negative eigenvector. If there are two negative eigenvalues, the stable manifold is a plane going through the steady state. This plane is generated by the two negative eigenvalues. In an $n \times n$ system, the stable arm (sometimes called the *stable manifold*) is a hyperplane generated by the associated eigenvectors, with dimension equal to the number of negative eigenvalues.

Analytical Solutions of Linear, Nonhomogeneous Systems Consider now the nonhomogeneous system of differential equations,

$$\dot{y}(t) = A \cdot y(t) + x(t) \tag{A.30}$$

where $y(t)$ is an $n \times 1$ vector of functions of time, $\dot{y}(t)$ is the corresponding vector of time derivatives, A is an $n \times n$ matrix of constants, and $x(t)$ is an $n \times 1$ vector of known functions of time, where these functions can be constants. The procedure to find the solutions to equation (A.30) parallels the one that we used for the homogeneous case. Begin again with the matrix V, composed of the eigenvectors of A, such that $V^{-1}AV$ generates a diagonal matrix D, which contains the eigenvalues of A. Transform the system by premultiplying all terms by V^{-1} and then define $z \equiv V^{-1}y$ to get

$$\dot{z} = V^{-1}\dot{y} = V^{-1} \cdot (Ay + x) = V^{-1}AVV^{-1}y + V^{-1}x = Dz + V^{-1}x$$

This matrix equation defines a system of n linear differential equations of the form

$$\dot{z}_i(t) = \alpha_i \cdot z_i(t) + V_i^{-1} \cdot x(t)$$

where V_i^{-1} is the ith row of V^{-1}. As we saw in section A.2.2, the solution to each of these linear ODEs with fixed coefficients takes the form of equation (A.11):

$$z_i(t) = e^{\alpha_i t} \cdot \int e^{-\alpha_i \tau} \cdot V_i^{-1} \cdot x(\tau) \cdot d\tau + e^{\alpha_i t} \cdot b_i$$

for $i = 1, \dots, n$, where b_i is again an arbitrary constant of integration. We can write these solutions in matrix notation as

$$z = E\hat{X} + Eb \tag{A.31}$$

where, again, E is a diagonal matrix of terms $e^{\alpha_i t}$, $\hat{X}$ is a column vector with integrals of the form

$$\int e^{-\alpha_i \tau} \cdot V_i^{-1} \cdot x(\tau) \cdot d\tau$$

as each of its elements, and b is a column vector of arbitrary constants. Once the time path of z is known, we can find the time path of y by premultiplying z by V.

As an example, consider the system of ODEs in equation (A.21). In matrix notation, this system can be written as

$$\begin{bmatrix} \dot{y}_1(t) \\ \dot{y}_2(t) \end{bmatrix} = \begin{bmatrix} 0.06 & -1 \\ -0.004 & 0 \end{bmatrix} \bullet \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} + \begin{bmatrix} 1.4 \\ 0.04 \end{bmatrix} \tag{A.32}$$

with the boundary conditions $y_1(0) = 1$ and

$$\lim_{t\to\infty} [e^{-0.06 \cdot t} \cdot y_1(t)] = 0$$

In this example, x is a vector of constants. In section A.5 we show how to find the eigenvalues and eigenvectors associated with a matrix A. We find that the diagonal matrix of eigenvalues, D, and the matrix of eigenvectors, V, are given by

$$D = \begin{bmatrix} 0.1 & 0 \\ 0 & -0.4 \end{bmatrix}, \qquad V = \begin{bmatrix} 1 & 1 \\ -0.04 & 0.1 \end{bmatrix}$$

where

$$V^{-1} = \begin{bmatrix} 0.1/0.14 & -1/0.14 \\ 0.04/0.14 & 1/0.14 \end{bmatrix}$$

Define $\begin{matrix} z_1 \\ z_2 \end{matrix} = V^{-1} \bullet \begin{matrix} y_1 \\ y_2 \end{matrix}$. The system in terms of the new variables can be written as

$$\dot{z}_1 = 0.1 \cdot z_1 + 10/14$$

$$\dot{z}_2 = -0.04 \cdot z_2 + 9.6/14$$

a system of two differential equations that we know how to solve (see section A.2.2):

$$z_1(t) = -100/14 + b_1 e^{0.1 \cdot t}$$

$$z_2(t) = 240/14 + b_2 e^{-0.04 \cdot t}$$

where b_1 and b_2 are constants of integration, which have to be pinned down by the boundary conditions. We can transform the solution for z_1 and z_2 into a solution for y_1 and y_2 by premultiplying z by V to get

$$y_1(t) = 10 + b_1 e^{0.1 \cdot t} + b_2 e^{-0.04 \cdot t} \tag{A.33}$$

$$y_2(t) = 2 - 0.04 \cdot b_1 e^{0.1 \cdot t} + 0.1 \cdot b_2 e^{-0.04 \cdot t} \tag{A.34}$$

We now need to determine the values of the constants, b_1 and b_2. The initial condition $y_1(0) = 1$ implies $b_1 + b_2 = -9$. We can multiply both sides of equation (A.33) by $e^{-0.06 \cdot t}$, take limits as t goes to infinity, and use the terminal condition, $\lim_{t \to \infty}[e^{-0.06t} \cdot y_1(t)] = 0$, to get

$$\lim_{t \to \infty} [e^{-0.06 \cdot t} \cdot y_1(t)] = \lim_{t \to \infty} [10 \cdot e^{-0.06 \cdot t} + b_1 e^{0.04 \cdot t} + b_2 e^{-0.1 \cdot t}] = 0$$

The first and third terms in the middle expression go to 0 as t goes to infinity, but the second term approaches infinity unless b_1 equals 0. Hence, the condition for the whole expression to equal 0 is $b_1 = 0$, which implies $b_2 = -9$. The exact solution to the system of ODEs is therefore

$$y_1(t) = 10 - 9 \cdot e^{-0.04 \cdot t}$$

$$y_2(t) = 2 - 0.9 \cdot e^{-0.04 \cdot t}$$

Note that $y_1(t)$ equals 1 at $t = 0$, increases over time, and asymptotes to its steady-state value, $y_1^* = 10$ (see figure A.10a). The variable y_2 equals 1.1 at $t = 0$, increases over time, and asymptotes to its steady-state value, $y_2^* = 2$ (see figure A.10b). In other words, the boundary conditions select the initial value of y_2 that causes the system to end up at its steady state. In terms of figure A.7a, the value $y_2(0)$ is chosen so as to put the system on the stable arm. At the initial point, $[\begin{smallmatrix} y_1(0) \\ y_2(0) \end{smallmatrix}] = [\begin{smallmatrix} 1 \\ 1.1 \end{smallmatrix}]$, the vector going toward the steady state is $[\begin{smallmatrix} 9 \\ 0.9 \end{smallmatrix}]$ or, by normalizing the first element to unity, $[\begin{smallmatrix} 1 \\ 0.1 \end{smallmatrix}]$, the negative eigenvector. Hence, as noted before,

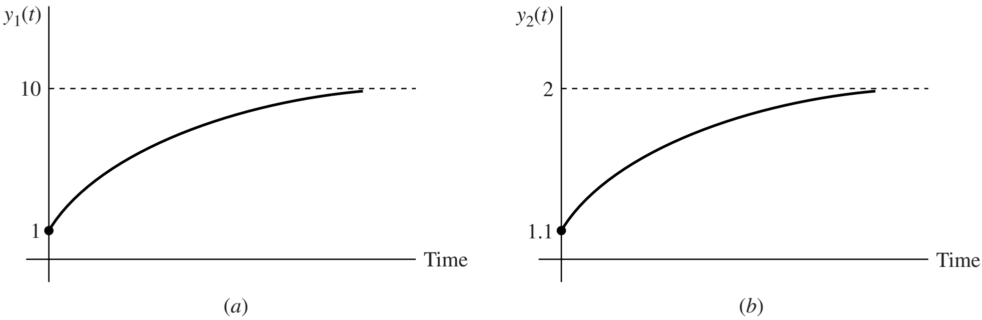


Figure A.10
(*a*) **Solution for $y_1(t)$.** The figure shows the solution for $y_1(t)$ in the system in equation (A.32). (*b*) **Solution for $y_2(t)$.** The figure shows the solution for $y_2(t)$ in the system in equation (A.32).

the stable arm goes through the steady state and corresponds to the eigenvector associated with the negative eigenvalue.

Linearization of Nonlinear Systems Many of the systems of ODEs that we encounter in the book are nonlinear. In this case, we can use the phase-diagram techniques that we discussed before, or alternatively, we can approximate the equations linearly by means of Taylor-series expansions.

Consider the following system of ODEs:

$$\begin{aligned} \dot{y}_1(t) &= f^1[y_1(t), \ldots, y_n(t)] \\ \dot{y}_2(t) &= f^2[y_1(t), \ldots, y_n(t)] \\ &\cdots \\ \dot{y}_n(t) &= f^n[y_1(t), \ldots, y_n(t)] \end{aligned} \tag{A.35}$$

where the functions $f^1(\bullet), f^2(\bullet), \ldots, f^n(\bullet)$ are nonlinear. We can use a Taylor-series expansion to study the system's dynamics in the neighborhood of its steady state. (Taylor's theorem is in section A.6.2.) The first-order expansion can be written as

$$\begin{aligned} \dot{y}_1(t) &= f^1(\bullet) + (f^1)_{y_1}(\bullet) \cdot (y_1 - y_1^*) + \cdots + (f^1)_{y_n}(\bullet) \cdot (y_n - y_n^*) + R_1 \\ &\cdots \\ \dot{y}_n(t) &= f^n(\bullet) + (f^n)_{y_1}(\bullet) \cdot (y_1 - y_1^*) + \cdots + (f^n)_{y_n}(\bullet) \cdot (y_n - y_n^*) + R_n \end{aligned} \tag{A.36}$$

where $f^1(\bullet), \ldots, f^n(\bullet)$ are the values of the functions $f^1(\bullet), \ldots, f^n(\bullet)$ at the steady state, and $(f^1)_{y_i}(\bullet), \ldots, (f^n)_{y_i}(\bullet)$ are the partial derivatives with respect to y_i at the steady state. The terms R_i are the Taylor residuals. If the system is close to its steady state, these residuals are small and can be neglected. The convenience of linearizing around the steady state is that, by definition of a steady state, the first element in each of the equations—$f^1(\bullet), \ldots, f^n(\bullet)$—is 0; that is, the steady-state value of $\dot{y}_i$ is zero for all i.

The linearized system in equation (A.36) can be written in matrix notation as

$$\dot{y} = A \cdot (y - y^*) \tag{A.37}$$

where A is an $n \times n$ matrix of constants corresponding to the first partial derivatives evaluated at the steady state. This linear system is similar to those analyzed in previous sections.

Consider the example of the system of nonlinear equations that we have already studied graphically,

$$\dot{k} = k^{0.3} - c \tag{A.22}$$

$$\dot{c} = c \cdot (0.3 \cdot k^{-0.7} - 0.06) \tag{A.23}$$

with the boundary conditions $k(0) = 1$ and $\lim_{t\to\infty}[e^{-0.06t} \cdot k(t)] = 0$. The steady-state values are $k^* = 10$ and $c^* = 2$. We can linearize this system as follows:

$$\begin{aligned} \dot{k} &= 0.3 \cdot (k^*)^{-0.7} \cdot (k - k^*) - (c - c^*) = 0.06 \cdot k - c + 1.4 \\ \dot{c} &= c^* \cdot [0.3 \cdot (-0.7) \cdot (k^*)^{-1.7}] \cdot (k - k^*) - 0 \cdot (c - c^*) = -0.008 \cdot k + 0.08 \end{aligned} \tag{A.38}$$

We know how to solve this linear system; in fact, we have already solved it! If we relabel k and c as y_1 and y_2, respectively, then it coincides with the system in equation (A.32).

As a graphical intuition, consider the phase diagram that we constructed for the nonlinear system defined by equations (A.22) and (A.23), as depicted in figure A.8. The loci in this figure are nonlinear. Around the steady state, however, the $\dot{c} = 0$ locus is vertical, and the $\dot{k} = 0$ locus is upward sloping. We can approximate these two loci with a vertical line and an upward-sloping line going through the same steady state. When the system is close to its steady state, this approximation is good. The approximation deteriorates as we move away from the steady state because the $\dot{k} = 0$ schedule is strictly concave. The dynamics of the nonlinear system is similar to that of the linear system in the vicinity of the steady state. In fact, at the steady state, the nonlinear stable arm corresponds to the negative eigenvector of the linearized system. Qualitatively, we see by comparing figures A.7a and A.8 that the two systems have similar dynamic properties.

The Time-Elimination Method for Nonlinear Systems In section A.2.3 we saw that one way to get a qualitative solution to a system of nonlinear differential equations was to use a phase diagram. The problem with this graphical approach is that it does not allow us to evaluate the model quantitatively. Later in that section we worked out an analytical solution to a linearized version of the system. The problem with this approach is that the quantification is local, valid only as an approximation in the neighborhood of the steady state. This section describes a method to find global numerical solutions to a system of ODEs. This method provides accurate results for a given configuration of parameters.

Consider again the system of nonlinear equations defined by equations (A.22) and (A.23):

$$\dot{k}(t) = k(t)^{0.3} - c(t) \tag{A.22}$$

$$\dot{c}(t) = c(t) \cdot [0.3 \cdot k(t)^{-0.7} - 0.06] \tag{A.23}$$

with the boundary conditions $k(0) = 1$ and $\lim_{t\to\infty}[e^{-0.06t} \cdot k(t)] = 0$. The phase diagram for this system is in figure A.8. If we knew the initial values, $c(0)$ and $k(0)$, then standard numerical methods for solving differential equations would allow us to solve out for the entire paths of c and k by integrating equations (A.22) and (A.23) with respect to time.[5]

The problem is that $c(0)$ is unknown. Instead, we are given the transversality condition, a restriction that forces the initial value of c to be on the stable arm. The challenge is to express this condition in terms of the required value of $c(0)$. The usual solution involves a method called *shooting*. Start with a guess about $c(0)$ and then work out the time paths implied by the differential equations (A.22) and (A.23). Then check whether the time paths approach the steady state and therefore satisfy the transversality condition. If the paths miss—as is almost sure to be true on the first try—then the system eventually diverges from the steady state. In this case, adjust the guess accordingly; reduce the conjectured value of $c(0)$ if the prior guess is too high, and vice versa. An approximation to the correct $c(0)$ can be found by iterating many times in this manner.

Mulligan and Sala-i-Martin (1991) worked out a much more efficient numerical technique called the *time-elimination method*. The key to this method is to eliminate time from the equations, just as we do when we construct a phase diagram. Recall that the stable arm shown in figure A.8 expresses c as a function of k. In dynamic programming this function is sometimes called the *policy function*. Imagine for a moment that we had a closed-form solution to this policy function, $c = c(k)$. In this case, we could use equation (A.22) to express $\dot{k}$ as a function of k: $\dot{k} = k^{0.3} - c(k)$. Since we know $k(0)$, we could use standard numerical methods to solve this first-order differential equation in k. Once we knew the path for k, we could determine the path for c (since we know the policy function, $c[k]$).

The time-elimination method provides a numerical technique for working out the policy function, $c = c(k)$. The trick is to note that the slope of this function is given by the

5. When the boundary conditions of a problem take the form of a set of values for all the variables at a single point in time, we call it an *initial-value problem*. For instance, the present problem would be an initial-value problem if we replaced the transversality condition, $\lim_{t\to\infty}[e^{-0.06t} \cdot k(t)] = 0$, with some value for $c(0)$. In contrast, for a *boundary-value problem,* the boundary conditions apply at different points in time. The present system is a *boundary-value problem* because we are given an initial condition, $k(0) = 1$, which applies at $t = 0$, and a terminal condition, $\lim_{t\to\infty}[e^{-0.06t} \cdot k(t)] = 0$, which applies at $t = \infty$. Initial-value problems are much easier to solve numerically.

ratio of $\dot{c}$ to $\dot{k}$:

$$dc/dk = c'(k) = \dot{c}/\dot{k} = \frac{c(k) \cdot [0.3 \cdot k^{-0.7} - 0.06]}{k^{0.3} - c(k)} \tag{A.39}$$

where we used the formulas for $\dot{k}$ and $\dot{c}$ from equations (A.22) and (A.23). Time does not appear in equation (A.39); hence, the name time-elimination method.

Note that equation (A.39) is a differential equation in c, where the derivative, dc/dk, is with respect to k rather than t. To solve this equation numerically by standard methods, we need one boundary condition; that is, we have to know one point, (c, k), that lies on the stable arm. Although we do not know the initial pair, $[c(0), k(0)]$, we know that the policy function goes through the steady state, (c^*, k^*). We can therefore start from this point and then solve equation (A.39) numerically to determine the rest of the policy function.[6] Note that, by eliminating time, we transformed a difficult *boundary-value problem* into a much easier *initial-value problem*.

Before we implement this method, there is one more problem that must be addressed. The slope of the policy function at the steady state is

$$c'(k^*) = (\dot{c})^*/(\dot{k})^* = 0/0$$

which is an indeterminate form. There are two ways to solve this problem. The first one uses l'Hôpital's rule for evaluating indeterminate forms (see section A.6.3). In this example, the application of l'Hôpital's rule yields

$$c'(k^*) = [c^* \cdot (-0.21) \cdot (k^*)^{-1.7}]/[0.3 \cdot (k^*)^{-0.7} - c'(k^*)]$$

which implies a quadratic equation in $c'(k)$:

$$[c'(k^*)]^2 - [0.3 \cdot (k^*)^{-0.7}] \cdot c'(k^*) - 0.21 \cdot c^* \cdot (k^*)^{-1.7} = 0$$

This equation has two solutions for $c'(k^*)$:

$$c'(k^*) = [0.3 \cdot (k^*)^{-0.7} - \{[0.3 \cdot (k^*)^{-0.7}]^2 + 4 \cdot (0.21) \cdot c^* \cdot (k^*)^{-1.7}\}^{1/2}]/2 \tag{A.40}$$

$$c'(k^*) = [0.3 \cdot (k^*)^{-0.7} + \{[0.3 \cdot (k^*)^{-0.7}]^2 + 4 \cdot (0.21) \cdot c^* \cdot (k^*)^{-1.7}\}^{1/2}]/2 \tag{A.41}$$

There are two solutions because there are two trajectories that go through the steady state: the stable arm and the unstable arm. The phase diagram in figure A.8 suggests that the stable arm

6. We might have considered starting from the steady state and going backward in time to solve the original system of two differential equations numerically. This idea does not work, however, because $\dot{k}$ and $\dot{c}$ are 0 at the steady state. Therefore, if we start at the steady state, we do not know how to move backward in time; that is, we cannot tell from where we came.

is upward sloping and the unstable arm is downward sloping. Since the slope of the stable arm at the steady state is positive, it must be given by the solution in equation (A.41).

The second way to compute the steady-state is to realize that, at the steady state, the policy function corresponds to the negative eigenvector. In other words, the slope of the negative eigenvector coincides with the steady-state slope of the policy function. Hence, we can use this value as the initial slope and then use equation (A.39) to compute the whole policy function. The advantage of the eigenvalue method over the l'Hôpital's rule method is that it does not require prior qualitative information about the sign of the steady-state slope.

The time-elimination method can be readily extended to systems of three differential equations with two controls and one state variable (see Mulligan and Sala-i-Martin, 1991, 1993). Consider a nonlinear system of equations,

$$\begin{aligned} \dot{c}(t) &= c[c(t), u(t), k(t)] \\ \dot{u}(t) &= u[c(t), u(t), k(t)] \\ \dot{k}(t) &= k[c(t), u(t), k(t)] \end{aligned} \tag{A.42}$$

where $c(t)$ and $u(t)$ are control variables, and $k(t)$ is the state variable. Imagine that we are given the initial value $k(0)$ and two transversality conditions (which apply at $t = \infty$). Suppose that the steady-state values are c^*, u^*, and k^*. Again, if we knew $c(0)$ and $u(0)$, we could find the solution to equation (A.42) by integrating with respect to time. The problem, however, is that $c(0)$ and $u(0)$ are unknown.

Imagine for the moment that we had closed-form expressions for $c(k)$ and $u(k)$, the policy functions for the problem. In this case we could plug these two functions into the $\dot{k}$ equation to get a single differential equation in k. Since we know $k(0)$, the whole time path for $k(t)$ could be found by integrating this differential equation with respect to time. Once we knew the path for k, we could determine the paths for c and u by plugging $k(t)$ into the two functions $c(k)$ and $u(k)$.

The time-elimination method provides a simple way to find $c(k)$ and $u(k)$ numerically. Use the chain rule of calculus to eliminate time from equation (A.42) to get the slopes of $c(k)$ and $u(k)$ as follows:

$$\begin{aligned} dc/dk = c'(k) = \dot{c}/\dot{k} &= \frac{c[c(k), u(k), k]}{k[c(k), u(k), k]} \\ du/dk = u'(k) = \dot{u}/\dot{k} &= \frac{u[c(k), u(k), k]}{k[c(k), u(k), k]} \end{aligned} \tag{A.43}$$

We can solve this system numerically by using the steady state, (c^*, u^*, k^*), as the initial condition. The steady-state slopes can be found by using l'Hôpital's rule or by computing the slope of the eigenvector associated with the negative eigenvalue.

A.2 Static Optimization

A.2.1 Unconstrained Maxima

Consider a univariate real function $u(\bullet)$. We say that a function $u(x)$ has a local maximum at $\overline{\overline{x}}$ if for all x in the neighborhood of $\overline{\overline{x}}$ (that is, for all x in the interval $[\overline{\overline{x}} - \epsilon, \overline{\overline{x}} + \epsilon]$, where ϵ is some positive number), $u(\overline{\overline{x}}) \geq u(x)$. We say that $u(x)$ has an absolute maximum[7] at $\overline{\overline{x}}$ if for all x in the domain of u, $u(\overline{\overline{x}}) \geq u(x)$.

Let $u(x)$ be twice continuously differentiable in the closed interval $[a, b]$ and let $\overline{\overline{x}}$ in the interior of $[a, b]$ be a local maximum. A *necessary condition* for $\overline{\overline{x}}$ to be an *interior local maximum* is for the first derivative of $u(\bullet)$ evaluated at $\overline{\overline{x}}$ to be 0, $u'(\overline{\overline{x}}) = 0$, and for the second derivative to be nonpositive, $u''(\overline{\overline{x}}) \leq 0$. If $u'(\overline{\overline{x}}) = 0$ and $u''(\overline{\overline{x}}) \leq 0$, then $\overline{\overline{x}}$ is an interior local maximum. That is, if the objective function is strictly concave (a negative second derivative), the necessary condition $u'(\overline{\overline{x}}) = 0$ is also a *sufficient condition*.

For practical purposes, if we want to find the maximum of a function in some interval, we compute the first derivative of that function and find the values of x that satisfy the equation $u'(\overline{\overline{x}}) = 0$. This condition gives us some candidate points, often called *critical points*. We then compute the second derivative of $u(\bullet)$ and evaluate it at the critical points. If it is negative, the critical point is a local maximum. We then compare the value $u(\overline{\overline{x}})$ with the value of the function at each of the corners a and b. The absolute maximum of $u(\bullet)$ in the interval $[a, b]$ occurs at one of the $\overline{\overline{x}}$, a, or b, depending on which has the largest image.

The multidimensional case is similar to the unidimensional case that we just described. Consider a function $u : R^n \to R$, twice continuously differentiable. A necessary condition for $u(x)$ to have an interior local maximum at $\overline{\overline{x}}$ (where x is now an n-dimensional vector, $x \equiv [x_1, \ldots, x_n]$) is for all the partial derivatives to vanish when evaluated at $\overline{\overline{x}}$. In other words, just as in the unidimensional case, functions are "flat at the top."

These necessary conditions are not sufficient, however, because local minima and saddle points also satisfy them. As a parallel to the unidimensional case, a sufficient condition is for the function u to be strictly concave at the critical point.[8]

7. A function $u(\bullet)$ achieves a minimum at point $\overline{\overline{x}}$ if $-u(\bullet)$ achieves a maximum at that point. Hence, to analyze minima of the function $u(\bullet)$, we can analyze maxima of $-u(\bullet)$.

8. One way to check strict concavity is to determine the definiteness of the Hessian, the matrix of second derivatives: if the Hessian is negative definite, the function u is strictly concave. A matrix is *negative definite* if and only if all its eigenvalues are negative. A matrix is *negative semidefinite* if and only if all its eigenvalues are nonpositive. A matrix is *positive definite* if and only if all its eigenvalues are positive. A matrix is *positive semidefinite* if and only if all its eigenvalues are nonnegative. A matrix is not definite if its eigenvalues do not all have the same signs. As we argued in a previous section, if we want to know the signs of the eigenvalues, we do not necessarily have to calculate them. For instance, in the 2×2 case, if the determinant of a matrix is negative, the eigenvalues must have opposite signs, because the determinant of the matrix equals the product of its eigenvalues.

A.2.2 Classical Nonlinear Programming: Equality Constraints

Suppose that we want to find the maximum of the function $u : R^n \to R$, subject to the constraint that the chosen point lie along a plane generated by the restriction $g(x) = a$, where $g : R^n \to R$, and x is an n-dimensional vector, $x \equiv (x_1, \ldots, x_n)$. That is, the problem is

$$\max_{x_1,\ldots,x_n} [u(x_1, \ldots, x_n)], \quad \text{subject to } g(x_1, \ldots, x_n) = a \tag{A.44}$$

We assume that $u(\bullet)$ and $g(\bullet)$ are twice continuously differentiable. One easy way to solve this problem is to realize that the restriction describes an implicit function for x_1: $x_1 = \tilde{x}_1(x_2, \ldots, x_n)$. (We assume here that the restriction uniquely determines x_1 for given values of $x_2, \ldots, x_n$.) We can plug the result for x_1 into $u(x)$ to get an unconstrained function of $(x_2, \ldots, x_n)$:

$$u[\tilde{x}_1(x_2, \ldots, x_n), (x_2, \ldots, x_n)] \equiv \tilde{u}(x_2, \ldots, x_n) \tag{A.45}$$

As just mentioned, the necessary condition for an unconstrained maximum of a function is for all the partial derivatives to vanish. When taking partial derivatives of $u(\bullet)$ with respect to each of the arguments x_i, $i = 2, \ldots, n$, we have to realize that $u(\bullet)$ depends on x_i directly and also indirectly through the dependence of x_1 on x_i. Hence, the necessary condition for a constrained maximum is

$$\partial\tilde{u}(\bullet)/\partial x_i = [\partial u(\bullet)/\partial x_1] \cdot \partial\tilde{x}_1/\partial x_i + \partial u(\bullet)/\partial x_i = 0 \tag{A.46}$$

for $i = 2, \ldots, n$. We can calculate the partial derivatives $\partial\tilde{x}_1/\partial x_i$ from the implicit function theorem (section A.6.1), $\partial\tilde{x}_1/\partial x_i = -[\partial g(\bullet)/\partial x_i]/[\partial g(\bullet)/\partial x_1]$. By plugging this expression into equation (A.46) we get

$$\frac{\partial g(\bullet)/\partial x_i}{\partial g(\bullet)/\partial x_1} = \frac{\partial u(\bullet)/\partial x_i}{\partial u(\bullet)/\partial x_1} \tag{A.47}$$

Another way to satisfy these conditions is for each of the partial derivatives of g with respect to x_i to be proportional to the partial derivative of u with respect to x_i, where the constant of proportionality μ is the same for all i. This set of conditions can be written in matrix notation as

$$Du(\overline{\overline{x}}) = \mu \cdot Dg(\overline{\overline{x}}) \tag{A.48}$$

where $\overline{\overline{x}}$ is an n-dimensional vector, and Dg and Du are the vectors of partial derivatives of g and u with respect to each of their arguments ($Dg \equiv [\partial g(\bullet)/\partial x_1, \ldots, \partial g(\bullet)/\partial x_n]$, and analogously for Du). The vectors Dg and Du are called the *gradients* of g and u,

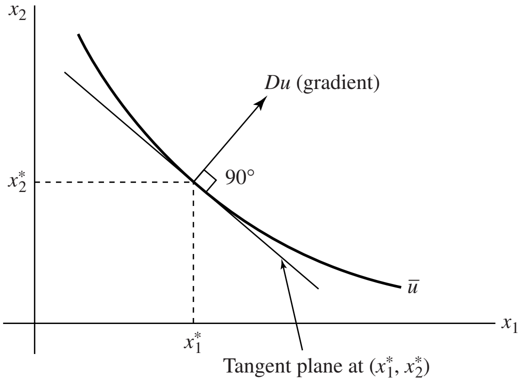


Figure A.11
Solution to a maximization problem subject to equality constraints. The figure illustrates the solution from equation (A.48), which involves a Lagrange multiplier, μ.

respectively. The gradient of a function $u(\bullet)$ evaluated at a point $\overline{\overline{x}}$ is a vector perpendicular to the tangent line of the function at that point (see figure A.11). Equation (A.48) says that a necessary condition for $\overline{\overline{x}}$ to be a maximum of the constrained problem is for the gradient of the restriction to be proportional to the gradient of the objective function at that point. The factor of proportionality is often called the *Lagrange multiplier,* μ. If we think of $u(\bullet)$ as a utility function and $g(\bullet) = a$ as a budget constraint (total spending, $g[\bullet]$, equals total income, a), then equation (A.48) is the familiar equality between marginal rates of substitution and marginal rates of transformation (or relative prices).

A convenient device for the derivation of these first-order conditions is the *Lagrangian,* which adds to the objective function a constant μ times the constraint:

$$L(\bullet) = u(x_1, \ldots, x_n) + \mu \cdot [a - g(x_1, \ldots, x_n)] \tag{A.49}$$

The first-order conditions in equation (A.48) are found by taking derivatives of the Lagrangian with respect to each of its arguments. Note that the derivative with respect to the Lagrange multiplier, μ, recovers the constraint.

To give an economic interpretation to the Lagrange multiplier, consider the change in utility, $u(\bullet)$, when income, a, changes. The total change in utility is given by

$$du(\bullet)/da = \sum_{i=1}^{n} [\partial u(\bullet)/\partial x_i] \cdot \partial \overline{\overline{x}}_i/\partial a$$

where $\partial \overline{\overline{x}}_i/\partial a$ is the change in the optimal quantity of good x_i when the constraint is relaxed by the amount ∂a. We can use the first-order conditions in equation (A.48) to rewrite this expression as

$$du(\bullet)/da = \sum_{i=1}^{n} \mu \cdot [\partial g(\bullet)/\partial x_i] \cdot \partial \overline{\overline{x}}_i/\partial a \tag{A.50}$$

If we totally differentiate the budget constraint with respect to a, we get

$$dg(\bullet)/da = \sum_{i=1}^{n} [\partial g(\bullet)/\partial x_i] \cdot \partial \overline{\overline{x}}_i/\partial a = 1$$

Substitution of this result into equation (A.50) implies

$$du(\bullet)/da = \mu \tag{A.51}$$

In other words, the Lagrange multiplier, μ, represents the extra utility that the agent gets when the constraint is relaxed by one unit. The Lagrange multiplier is therefore often referred to as the *shadow price* or *shadow value of the constraint.* This interpretation is important and will be used throughout the book.

A.2.3 Inequality Constraints: The Kuhn–Tucker Conditions

Imagine now that an agent faces m inequality restrictions of the form

$$g_i(x_1, \ldots, x_n) \le a_i \quad \text{for } i = 1, \ldots, m$$

All the functions $g_i(\bullet)$ are assumed to be twice continuously differentiable, and each a_i is constant. The problem can be written as

$$\begin{aligned} &\max_{x_1,\ldots,x_n} [u(x_1, \ldots, x_n)], \quad \text{subject to} \\ &\qquad g_1(x_1, \ldots, x_n) \le a_1 \\ &\qquad\qquad \ldots \\ &\qquad g_m(x_1, \ldots, x_n) \le a_m \end{aligned} \tag{A.52}$$

Most economic constraints take the form shown in equation (A.52). For example, a budget constraint does not require an agent to spend all of his income but says that he cannot spend more than his income.

An easy way to solve the problem in equation (A.52) is to use the Kuhn–Tucker (1951) theorem. The theorem says that if $\overline{\overline{x}} = (\overline{\overline{x}}_1, \ldots, \overline{\overline{x}}_n)$ is a solution to problem (A.52),[9] then

9. An additional condition is that the “constraint qualification” be satisfied. This condition requires the gradients of the constraints to be linearly independent.

there exists a set of m Lagrange multipliers such that

$$\begin{aligned}
&(a)\ \ Du(\bullet) = \sum_{i=1}^{m} \mu_i \cdot [Dg_i(\bullet)] \\
&(b)\ \ g_i(\bullet) \leq a_i,\ \mu_i \geq 0 \\
&(c)\ \ \mu_i \cdot [a_i - g_i(\bullet)] = 0
\end{aligned} \tag{A.53}$$

Condition (a) in equation (A.53) says that the gradient of the objective function must be a linear combination of the gradients of the restrictions. The weights in this linear combination are the Lagrange multipliers. In the particular case when there is only one restriction, $m = 1$, this condition is equivalent to equation (A.48). Condition (b) in equation (A.53) says that for $\overline{\overline{x}}$ to be an optimum, the constraints have to be satisfied and the shadow prices must be nonnegative. That is, $Du(\bullet)$ must lie on the cone generated by the $Dg_i(\bullet)$.

Condition (c) in equation (A.53) is often called the *complementary-slackness condition.* It says that the product of the shadow price and the constraint is 0. This condition means that if the constraint $g_i(\bullet) - a_i$ is not binding (if it is not satisfied with strict equality), the shadow price must be 0. That is, $Dg_i(\bullet)$ receives no weight in the linear combination that generates $Du(\bullet)$. In contrast, if the price is strictly positive, the constraint associated with it must be binding.[10]

Consider the example in figure A.12. There are two constraints, $g_1(\bullet) \leq a_1$ and $g_2(\bullet) \leq a_2$. The first constraint restricts the set of points in the space to lie between the curve labeled g_1 and the origin. Similarly, the second constraint restricts to the space between the curve labeled g_2 and the origin. The objective function can be represented by a set of indifference curves labeled u_i, which increase in the northeast direction. The gradients of the two constraints (which point in the direction perpendicular to the tangent at that particular point) are labeled Dg_1 and Dg_2. Condition (a) says that if $\overline{\overline{x}}$ is to be an optimum, the gradient of $u(\bullet)$ must be a linear combination of the two gradients Dg_1 and Dg_2. Condition (b) says that the linear combination must involve nonnegative weights. Graphically, these conditions mean that the gradient of u must lie on the cone described by the gradients of the two constraints.

To understand the meaning of the complementary-slackness condition, imagine that the preferences for a pair of goods take the form of a bell (figure A.13a). The indifference curves are circles around a point that yields maximum utility. (This point would correspond to a level of satiation beyond which agents would not like to go, no matter what the prices are.) Suppose that the budget constraint lies to the left of this satiation point (see figure A.13b).

10. In economic terms, the complementary-slackness condition says that if a constraint is not binding (that is, if it is unimportant) and we relax it by one unit, the attained utility does not change.

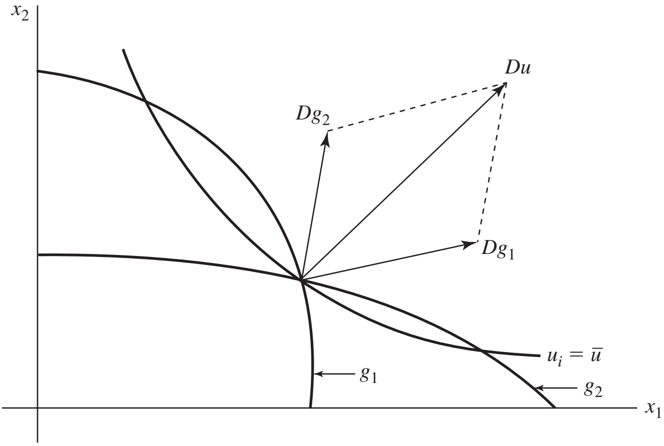


Figure A.12
Solution to a maximization problem subject to inequality constraints. The figure illustrates the solution to a maximization problem of the form of equation (A.53) with two inequality constraints.

The agent would like to consume more of both goods, but the budget constraint does not permit him or her to do so. Hence, the constraint is binding. The Kuhn–Tucker theorem says that the gradient of the objective function at the optimum is proportional to the gradient of the constraint. Since the gradient is perpendicular to the function, this condition means that the maximum occurs at the tangency point.

Consider now what happens when the satiation point is fully inside the budget set (figure A.13c). The individual clearly achieves maximum utility by remaining inside the budget set. In other words, since the constraint is not binding, the agent behaves as if he were not constrained. The Kuhn–Tucker theorem says that, at the optimum, the gradient of the objective function is proportional to the gradient of the constraint. The complementary-slackness condition says that when the constraint is not binding, the factor of proportionality is 0. Hence, the gradient of the objective function must equal 0, the condition for an unconstrained maximum. To summarize, the complementary-slackness condition says that if a constraint is not binding, it will not affect the optimal choice.

The Kuhn–Tucker conditions can be read another way by writing the Lagrangian function as

$$L(x_1, \ldots, x_n; \mu_1, \ldots, \mu_m) = u(x_1, \ldots, x_n) + \sum_{i=1}^{m} \mu_i \cdot [a_i - g_i(x_1, \ldots, x_n)] \tag{A.54}$$

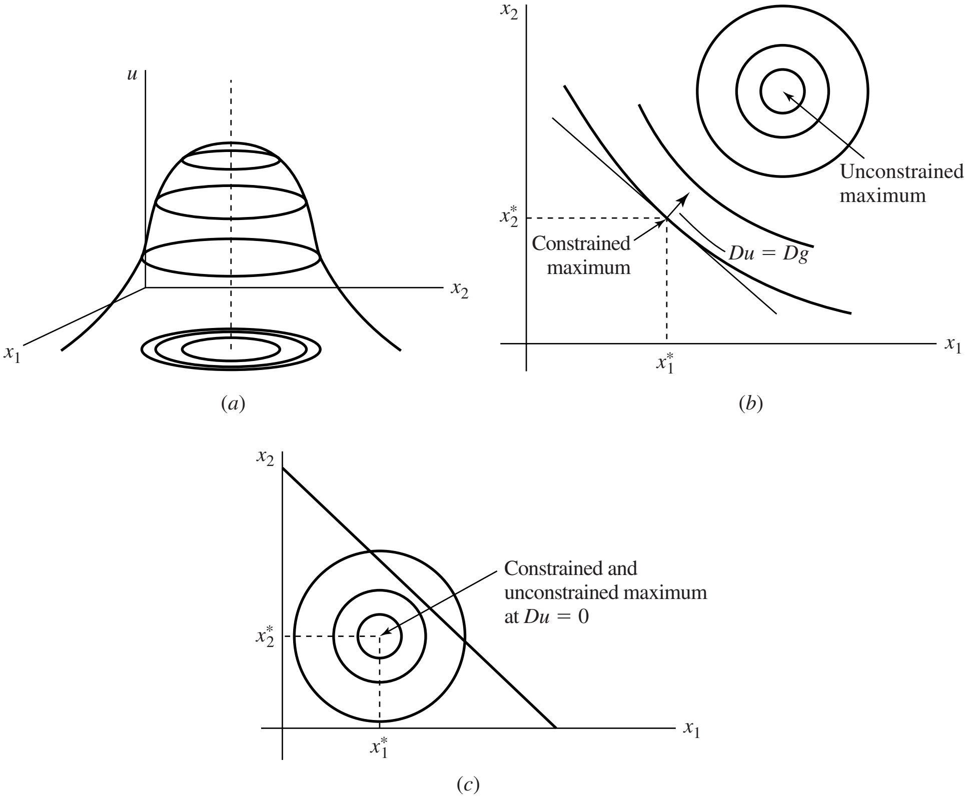


Figure A.13
(a) **Preferences over two goods.** The indifference curves for x_1 and x_2 are assumed to take the form of a bell. (b) **Maximizing utility subject to a binding inequality constraint.** In this example, the budget constraint for x_1 and x_2 is binding. (c) **Maximizing utility subject to a nonbinding inequality constraint.** In this example, the budget constraint for x_1 and x_2 is not binding.

Condition (a) in equation (A.53) says that a necessary condition for the vector $\overline{\overline{x}}$ to be a maximum of the constrained problem is for it to be a maximum of the associated Lagrangian. Conditions (b) and (c) in equation (A.53) say that, at the optimum, the Lagrangian has a minimum with respect to the vector $\mu \equiv (\mu_1, \ldots, \mu_m)$. (Condition [$b$] says that the two components in [c] are nonnegative; hence, the product of the two is minimized at 0.) Taken together, conditions (a)–(c) in equation (A.53) say that a necessary condition for $\overline{\overline{x}}$ to be an optimum is for the Lagrangian to have a saddle point at $(\overline{\overline{x}}, \mu)$; that is, a maximum with respect to x and a minimum with respect to μ.

Conditions (a)–(c) in equation (A.53) are the set of necessary conditions from the Kuhn–Tucker theorem; if a point is to be an optimum, it must satisfy them. If the objective function $u(\bullet)$ is concave and the constraints form a convex set, the necessary conditions are also sufficient.[11]

A.3 Dynamic Optimization in Continuous Time

A.3.1 Introduction

Mathematicians have long worried about dynamic problems. It is commonly thought that the first person to solve one of these problems was Bernoulli in 1696. Euler and Lagrange also worked with dynamic problems. Most applications of their theoretical findings were in physics, especially as related to Hamilton's principle or the principle of least action. Economists have been interested in dynamic problems since at least the work of Hotelling and Ramsey in the 1920s. It was not until the 1960s, however, that dynamic mathematical techniques were widely introduced into economics, mainly in the work of the neoclassical growth theorists. These techniques are now part of the toolbox of most modern economists.

The methodology that classical mathematicians used to solve dynamic problems is known as the *calculus of variations.* This approach has since been generalized in two ways. First, Richard Bellman, an American mathematician, developed the method of *dynamic programming* in the 1950s. This method is especially suited to discrete-time problems and is particularly useful for stochastic models. Second, also in the 1950s, a team of Russian mathematicians led by L. Pontryagin developed the *maximum principle of optimal control.* (The first English translation of this work did not appear, however, until 1962.)

In this chapter, we demonstrate how to use Pontryagin's technique. The maximum principle is a generalization of the classical calculus of variations in that it provides solutions to problems in which one or more of the constraints involve the derivatives of some of the state variables. This type of constraint is central to the theory of economic growth.

Our goal in this section is not to prove the maximum principle but, rather, to provide a heuristic derivation along with a description of the procedure that we follow to use the solutions. This approach will provide us with a set of tools that will allow us to solve the various dynamic models that will be encountered in the book.[12]

11. A slightly less restrictive set of sufficient conditions is given by Arrow and Enthoven (1961): they require the objective function to be quasi-concave, that is, to exhibit convex upper-level sets.

12. A full proof of the maximum principle is in Pontryagin et al. (1962).

A.3.2 The Typical Problem

The typical problem that we want to solve takes the following form. The agent chooses or controls a number of variables, called *control variables,*[13] so as to maximize an objective function subject to some constraints. These constraints are dynamic in that they describe the evolution of the state of the economy, as represented by a set of *state variables,* over time. The problem is given by

$$\max_{c(t)} V(0) = \int_0^T v[k(t), c(t), t] \cdot dt, \quad \text{subject to}$$

$$(a)\;\; \dot{k}(t) = g[k(t), c(t), t] \tag{A.55}$$

$$(b)\;\; k(0) = k_0 > 0$$

$$(c)\;\; k(T) \cdot e^{-\bar{r}(T)\cdot T} \geq 0$$

where $V(0)$ is the value of the objective function as seen from the initial moment 0, $\bar{r}(t)$ is an average discount rate that applies between dates 0 and t, and T is the terminal planning date, which could be finite or infinite. We discuss the difference between a finite and an infinite horizon in section A.4.7.

The variable $k(t)$—which appears with an overdot in part (a) of equation (A.55)—is the *state variable,* and the variable $c(t)$ is the *control variable*. Each of these variables is a function of time. The objective function in equation (A.55) is the integral of instantaneous felicity functions, $v(\bullet)$,[14] over the interval from 0 to T. These felicity functions depend on the state and control variables, $k(t)$ and $c(t)$, and on time, t.

The accumulation constraint in part (a) of equation (A.55) is a differential equation in $k(t)$; this constraint shows how the choice of the control variable, $c(t)$, translates into a pattern of movement for the state variable, $k(t)$. The expression for $\dot{k}(t)$ is called the *transition equation* or *equation of motion*. Although we have only one transition equation, there is a continuum of constraints, one for every point in time between 0 and T.[15]

The initial condition in part (b) of equation (A.55) says that the state variable, $k(t)$, begins at a given value, k_0. The final constraint, in part (c) of equation (A.55), says that the chosen

13. Pontryagin et al. (1962) call these control variables *steering variables*.

14. Examples of felicity functions are utility functions of consumers, profit functions of firms, and objective functions of governments. To fix ideas, in this chapter we identify them with utility functions.

15. This accumulation equation could be cast as an inequality restriction, $\dot{k} \leq g(\bullet)$. Typically, individuals will not find it optimal to satisfy this restriction with strict inequality because it will be advantageous to increase $c(t)$ to raise the current flow of utility or to increase $k(t)$ to raise the future flows of utility. We therefore leave the restriction as an equality.

value of the state variable at the end of the planning horizon, $k(T)$, discounted at the rate $\bar{r}(T)$, must be nonnegative. For finite values of T, this constraint implies $k(T) \geq 0$, as long as the discount rate $\bar{r}(T)$ is positive and finite. If $k(t)$ represents a person's net assets and T the person's lifetime, the constraint in part (c) of equation (A.55) precludes dying in debt. If the planning horizon is infinite, the condition says that net assets can be negative and grow forever in magnitude, as long as the rate of growth is less than $\bar{r}(t)$. This constraint rules out chain letters or Ponzi schemes for debt.

An economic example of a dynamic problem of this kind is a growth model in which $v(\bullet)$ is an instantaneous utility function that depends on the level of consumption and is discounted by a time-preference factor,

$$v(k, c, t) = e^{-\rho t} \cdot u[c(t)] \tag{A.56}$$

In this example, $v(\bullet)$ does not depend on the capital stock, $k(t)$, and depends directly on time only through the discount factor, $e^{-\rho t}$. The constraint describes the accumulation of the variable $k(t)$. If we think of $k(t)$ as physical capital, an example of such a constraint is

$$\dot{k} = g[k(t), c(t), t] = f[k(t), t] - c(t) - \delta \cdot k(t) \tag{A.57}$$

where δ is the fraction of the capital stock that depreciates at every instant. Equation (A.57) says that the increase in the capital stock (net investment) equals total saving minus depreciation. Total saving, in turn, equals the difference between output, $f(\bullet)$, and consumption, $c(t)$. The dependence of production on t, for given $k(t)$, could reflect the state of technology or knowledge at a given point in time.

A.3.3 Heuristic Derivation of the First-Order Conditions

A formal proof of the maximum principle is outside the scope of this book; we will instead provide a heuristic derivation. Readers who are interested only in the procedure for finding the first-order conditions, and not in the derivation, can skip sections A.4.3–A.4.9 and go directly to section A.4.10.

The starting point is the static method for solving nonlinear optimization problems, the Kuhn–Tucker Theorem. This theorem, described in section A.3.3, suggests the construction of a Lagrangian of the form,

$$L = \int_0^T v[k(t), c(t), t] \cdot dt + \int_0^T \{\mu(t) \cdot (g[k(t), c(t), t] - \dot{k}(t))\} \cdot dt + \nu \cdot k(T) \cdot e^{\bar{r}(T) \cdot T} \tag{A.58}$$

where $\mu(t)$ is the Lagrange multiplier associated with the constraint in part (a) of equation (A.55), and ν is the multiplier associated with the constraint in part (c) of equation (A.55).[16] Since there is a continuum of constraints from part (a), one for each instant t between 0 and T, there is a corresponding continuum of Lagrange multipliers, $\mu(t)$. The $\mu(t)$ are called *costate variables* or *dynamic Lagrange multipliers*. Following the parallel with the static case, these costate variables can be interpreted as shadow prices: $\mu(t)$ is the price or value of an extra unit of capital stock at time t in units of utility at time 0. Since each of the constraints, $g(\bullet) - \dot{k}$, equals 0, each of the products, $\mu(t) \cdot [g(\bullet) - \dot{k}]$, also equals 0. It follows that the "sum" of all of the constraints equals 0:

$$\int_0^T \{\mu(t) \cdot (g[k(t), c(t), t] - \dot{k}(t))\} \cdot dt = 0$$

This expression appears in the middle of equation (A.58).

To find the set of first-order necessary conditions in a static problem, we would maximize L with respect to $c(t)$ and $k(t)$ for all t between 0 and T. The problem with this procedure is that we do not know how to take the derivative of $\dot{k}$ with respect to k. To avoid this problem, we can rewrite the Lagrangian by integrating the term $\mu(t) \cdot \dot{k}(t)$ by parts to get[17]

$$\begin{aligned} L = &\int_0^T (v[k(t), c(t), t] + \mu(t) \cdot g[k(t), c(t), t])\, dt \\ &+ \int_0^T \mu(t) k(t)\, dt + \mu(0) k_0 - \mu(T) k(T) + \nu k(T) e^{-\bar{r}(T)T} \end{aligned} \tag{A.59}$$

The expression inside the first integral is called the *Hamiltonian* function,

$$H(k, c, t, \mu) \equiv v(k, c, t) + \mu \cdot g(k, c, t) \tag{A.60}$$

The Hamiltonian function has an economic interpretation (see Dorfman, 1969). At an instant in time, the agent consumes $c(t)$ and owns a stock of capital $k(t)$. These two variables affect utility through two channels. First, the direct contribution of consumption, and perhaps capital, to utility, is captured by the term $v(\bullet)$ in equation (A.60). Second, the choice of consumption affects the change in the capital stock in accordance with the transition equation for $\dot{k}$ in part (a) of equation (A.55). The value of this change in the capital stock

16. We would also have the constraints $c(t) \geq 0$, but commonly assumed forms of the utility function imply that these constraints will not be binding. We therefore ignore these inequality restrictions in the present discussion.

17. To integrate $\int_0^T (\dot{k}) \cdot \mu\, dt$ by parts, start with $(d/dt)(\mu k) = \dot{\mu} k + \dot{k} \mu$. Integrate both sides of this expression between 0 and T and note that $\int_0^T (d/dt)(k\mu) \cdot dt = k(T) \cdot \mu(T) - k(0) \cdot \mu(0)$. From this expression, subtract the integral of $k\dot{\mu}$ to get $\int_0^T (\dot{k}) \cdot \mu\, dt = k(T) \cdot \mu(T) - k(0) \cdot \mu(0) - \int_0^T (\dot{\mu}) \cdot k\, dt$, which is the expression used to compute equation (A.59). See sections A.6.4 and A.6.5 for further discussion.

is the term $\mu \cdot g(\bullet)$ in equation (A.60). Hence, for a given value of the shadow price, μ, the Hamiltonian captures the total contribution to utility from the choice of $c(t)$.

Rewrite the Lagrangian from equation (A.59) as

$$L = \int_0^T \{H[k(t), c(t), t] + \dot{\mu}(t) \cdot k(t)\} \cdot dt + \mu(0) \cdot k_0 - \mu(T) \cdot k(T) + \nu \cdot k(T) \cdot e^{-\bar{r}(T) \cdot T} \tag{A.61}$$

Let $\overline{\overline{c}}(t)$ and $\overline{\overline{k}}(t)$ be the optimal time paths for the control and state variables, respectively. If we perturb the optimal path $\overline{\overline{c}}(t)$ by an arbitrary perturbation function, $p_1(t)$, we can generate a neighboring path for the control variable,

$$c(t) = \overline{\overline{c}}(t) + \epsilon \cdot p_1(t)$$

When $c(t)$ is thus perturbed, there must be a corresponding perturbation to $k(t)$ and $k(T)$ so as to satisfy the budget constraint:

$$k(t) = \overline{\overline{k}}(t) + \epsilon \cdot p_2(t)$$

$$k(T) = \overline{\overline{k}}(T) + \epsilon \cdot dk(T)$$

If the initial paths are optimal, then $\partial L/\partial \epsilon$ should equal 0. Before we compute such a derivative, it will be convenient to rewrite the Lagrangian in terms of ϵ:

$$\overline{\overline{L}}(\cdot, \epsilon) = \int_0^T \{H[k(\bullet, \epsilon); c(\bullet, \epsilon)] + \dot{\mu}(\bullet) \cdot k(\bullet, \epsilon)\} \cdot dt$$
$$+ \mu(0) \cdot k_0 - \mu(T) \cdot k(T, \epsilon) + \nu \cdot k(T, \epsilon) \cdot e^{-\bar{r}(T) \cdot T}$$

We can now take the derivative of the Lagrangian with respect to ϵ and set it to 0:

$$\partial \overline{\overline{L}}/\partial \epsilon = \int_0^T [\partial H/\partial \epsilon + \dot{\mu} \cdot \partial k/\partial \epsilon] \cdot dt + [\nu e^{-\bar{r}(T)T} - \mu(T)] \cdot \partial k(T, \epsilon)/\partial \epsilon = 0$$

The chain rule of calculus implies $\partial H/\partial \epsilon = [\partial H/\partial c] \cdot p_1(t) + [\partial H/\partial k] \cdot p_2(t)$ and $\partial k(T, \epsilon)/\partial \epsilon = dk(T)$. Use these formulas and rearrange terms in the expression for $\partial \overline{\overline{L}}/\partial \epsilon$ to get

$$\partial L/\partial \epsilon = \int_0^T \{[\partial H/\partial c] \cdot p_1(t) + [\partial H/\partial k + \dot{\mu}] \cdot p_2(t)\} \cdot dt$$
$$+ [\nu \cdot e^{-\bar{r}(T)T} - \mu(T)] \cdot dk(T) = 0 \tag{A.62}$$

Equation (A.62) can hold for all perturbation paths, described by $p_1(t)$, $p_2(t)$, and $dk(T)$, only if each of the components in the equation vanishes, that is,

$$\partial H/\partial c = 0 \tag{A.63}$$

$$\partial H/\partial k + \dot{\mu} = 0 \tag{A.64}$$

$$\nu \cdot e^{-\bar{r}(T)\cdot T} = \mu(T) \tag{A.65}$$

The first-order condition with respect to the control variable in equation (A.63) says that if $\bar{\bar{c}}(t)$ and $\bar{\bar{k}}(t)$ are a solution to the dynamic problem, the derivative of the Hamiltonian with respect to the control c equals 0 for all t. This result is called the *maximum principle*. Equation (A.64) says that the partial derivative of the Hamiltonian with respect to the state variable equals the negative of the derivative of the multiplier, $-\dot{\mu}$. This result and the transition equation in part (a) of equation (A.55) are often called the Euler equations. Finally, equation (A.65) says that the costate variable at the terminal date, μ, equals ν, the static Lagrange multiplier associated with the nonnegativity constraint on k at the terminal date, discounted at the rate $\bar{r}(T)$.

A.3.4 Transversality Conditions

Section A.3.3 showed that the Kuhn–Tucker necessary first-order conditions include a complementary-slackness condition associated with the inequality constraints. In the static problem, these conditions say that if a restriction is not binding, the shadow price associated with it is 0. In the present dynamic problem, there is an inequality constraint that says that the stock of capital left at the end of the planning period, discounted at the rate $\bar{r}(T)$, cannot be negative, $k(T)\cdot e^{-\bar{r}(T)\cdot T} \geq 0$. The condition associated with this constraint requires $\nu \cdot k(T) \cdot e^{-\bar{r}(T)\cdot T} = 0$, with $\nu \geq 0$. Equation (A.65) implies that we can rewrite this complementary-slackness condition as

$$\mu(T)\cdot k(T) = 0 \tag{A.66}$$

This boundary condition is often called the *transversality condition*. It says that if the quantity of capital left is positive, $k(T) > 0$, its price must be 0, $\mu(T) = 0$. Alternatively, if capital at the terminal date has a positive value, $\mu(T) > 0$, the agent must leave no capital, $k(T) = 0$. We discuss later the meaning of equation (A.66) when T is infinite.

A.3.5 The Behavior of the Hamiltonian over Time

To see how the optimal value of the Hamiltonian behaves over time, take the total derivative of H with respect to t to get

$$dH(k, c, \mu, t)/dt = [\partial H/\partial k]\cdot \dot{k} + [\partial H/\partial c \cdot \dot{c}] + [\partial H/\partial \mu]\cdot \dot{\mu} + \partial H/\partial t \tag{A.67}$$

The first-order condition in equation (A.63) implies that, at the optimum, $\partial H/\partial c = 0$; hence, the second term on the right-hand side of equation (A.67) equals 0. Equation (A.64) requires $\partial H/\partial k = -\dot{\mu}$. Since $\partial H/\partial \mu = g = \dot{k}$, the first and third terms on the right-hand side of equation (A.67) cancel. Hence, at the optimum, the total derivative of the Hamiltonian with respect to time equals the partial derivative, $\partial H/\partial t$. If the problem is autonomous—that is, if neither the objective function nor the constraints depend directly on time—then the derivative of the Hamiltonian with respect to time is 0. In other words, the Hamiltonian associated with autonomous problems is constant at all points in time. These results on the behavior of the Hamiltonian will be used later in this appendix.

A.3.6 Sufficient Conditions

In a static, nonlinear maximization problem, the Kuhn–Tucker necessary conditions are also sufficient when the objective function is concave and the restrictions generate a convex set. Mangasarian (1966) extends this result to dynamic problems and shows that if the functions $v(\bullet)$ and $g(\bullet)$ in equation (A.55) are both concave in k and c, then the necessary conditions are also sufficient. This sufficiency result is easy to use but is somewhat restrictive.

More general sufficiency conditions are given by Arrow and Kurz (1970). Define $H^0(k, \mu, t)$ to be the maximum of $H(k, c, \mu, t)$ with respect to c, given k, μ, and t. The Arrow–Kurz theorem says that if $H^0(k, \mu, t)$ is concave in k, for given μ and t, then the necessary conditions are also sufficient. Concavity of $v(\bullet)$ and $g(\bullet)$ is sufficient, but not necessary, for the Arrow–Kurz condition to be satisfied. The disadvantage of this more general result is that checking the properties of a derived function, such as H^0, tends to be harder than checking the properties of $v(\bullet)$ and $g(\bullet)$.

A.3.7 Infinite Horizons

Most of the growth models that we discuss in the book involve economic agents with infinite planning horizons. The typical problem takes the form

$$\max_{c(t)} V(0) = \int_0^{\infty} v[k(t), c(t), t] \cdot dt, \quad \text{subject to}$$

$$\begin{aligned}
&(a)\ \ \dot{k}(t) = g[k(t), c(t), t] \\
&(b)\ \ k(t) = k_0 \\
&(c)\ \ \lim_{t\to\infty} [k(t) \cdot e^{-\bar{r}(t)\cdot t}] \geq 0
\end{aligned} \tag{A.68}$$

The only difference between equation (A.68) and equation (A.55) is that the planning horizon—the number on top of the integral—in equation (A.68) is infinity, rather than

$T < \infty$. The first-order conditions for the infinite-horizon problem are the same as those for the finite horizon case, equations (A.63) and (A.64). The key difference is that the transversality condition, shown in equation (A.66), applies not to a finite T, but to the limit as T tends to infinity. In other words, the transversality condition is now

$$\lim_{t\to\infty} [\mu(t) \cdot k(t)] = 0 \tag{A.69}$$

The intuitive explanation for the new condition is that the value of the capital stock must be asymptotically 0; otherwise, something valuable would be left over. If the quantity $k(t)$ remains positive asymptotically, then the price, $\mu(t)$, must approach 0 asymptotically. If $k(t)$ grows forever at a positive rate—as occurs in some of the models that we study in this book—then the price $\mu(t)$ must approach 0 at a faster rate so that the product, $\mu(t) \cdot k(t)$, goes to 0.

Although equation (A.69) has intuitive appeal as the limiting version of equation (A.66), there is disagreement over the conditions under which equation (A.69) is actually a necessary condition for the infinite-horizon problem in equation (A.68). Recall that the only argument we gave for its validity was the analogue to the transversality condition in the finite-horizon case. Some researchers have found counterexamples in which equation (A.69) is not a necessary condition for optimization. In section A.4.9 we discuss one of these examples.

One transversality condition that always applies was found by Michel (1982). He argues that the transversality condition requires the value of the Hamiltonian to approach 0 as t goes to infinity:

$$\lim_{t\to\infty} [H(t)] = 0 \tag{A.70}$$

We can derive this transversality condition if we follow Michel and think of the infinite-horizon case as a setting in which the agent chooses the terminal date, T. If we perturb the terminal date T in equation (A.61) by $\epsilon \cdot dT$, we find that the limit of integration now depends on ϵ. When we take derivatives of the Lagrangian with respect to ϵ, we find that one of the terms in equation (A.62) is $H(T) \cdot dT$. This term comes from taking the derivative of the limit of integration, $T(\epsilon)$, with respect to ϵ. As with all the terms in equation (A.62), this one will have to be 0 at the optimum. If the terminal date is fixed, so that $dT = 0$, then $H(T)$ can take on any value. But if the terminal date is variable, so that dT is nonzero, then $H(T)$ must vanish. If we take the limit as T goes to infinity, we get the transversality condition in equation (A.70). This condition is redundant in most of the models that we study in the book because it will be satisfied whenever equation (A.69) is satisfied. Thus, in most cases, we can use equation (A.69) and ignore equation (A.70).

A.3.8 Example: The Neoclassical Growth Model

We consider here the example of the neoclassical growth model with a Cobb–Douglas production function. (See chapter 2 for more details.) Assume that economic agents choose the path of consumption, $c(t)$, and capital, $k(t)$, so as to maximize the objective function,

$$U(0) = \int_0^{\infty} e^{-\rho t} \cdot \log[c(t)] \cdot dt$$

$$\begin{aligned} &(a)\;\; \dot{k}(t) = [k(t)^{\alpha} - c(t) - \delta \cdot k(t)] \\ &(b)\;\; k(0) = 1 \\ &(c)\;\; \lim_{t\to\infty} [k(t) \cdot e^{-\bar{r}(t)\cdot t}] \geq 0 \end{aligned} \tag{A.71}$$

where α is a constant with $0 < \alpha < 1$. We normalize the initial capital $k(0)$ to unity. The interest rate, $r(t)$, equals the net marginal product of capital, $\alpha \cdot k(t)^{\alpha-1} - \delta$, and the average interest rate, $\bar{r}(t)$, equals $(1/t) \cdot \int_0^t r(v) \cdot dv$.

The agent can be thought of as a household-producer who wants to maximize utility, represented as the present discounted value of a stream of instantaneous felicities. Each of these felicities depends on the instantaneous flow of consumption. The felicity function is assumed in equation (A.71) to be logarithmic. The discount rate is $\rho > 0$. The agent has access to the technology (the Cobb–Douglas form described in chapter 1) that transforms capital into output according to $y(t) = [k(t)]^{\alpha}$. The accumulation constraint in part (a) of equation (A.71) says that total output has to be divided between consumption, $c(t)$, depreciation, $\delta \cdot k(t)$, and capital accumulation, $\dot{k}(t)$. The initial condition in part (b) of equation (A.71) says that the capital stock at time 0 is 1. The restriction in part (c) of equation (A.71) says that the capital stock left over at the "end of the planning horizon," when discounted at the average interest rate, $\bar{r}(t)$, is nonnegative. (If $k[t]$ represents household assets, this condition precludes chain-letter policies in which debt accumulates forever at a rate at least as high as the interest rate.)

To solve the optimization problem, set up the Hamiltonian,

$$H(c, k, t, \mu) = e^{-\rho t} \cdot \log(c) + \mu \cdot (k^{\alpha} - c - \delta k) \tag{A.72}$$

Equations (A.63) and (A.64) imply that the first-order conditions are

$$H_c = e^{-\rho t} \cdot (1/c) - \mu = 0 \tag{A.73}$$

$$H_k = \mu \cdot (\alpha k^{\alpha-1} - \delta) = -\dot{\mu} \tag{A.74}$$

and equation (A.69) implies that the transversality condition is

$$\lim_{t\to\infty} [\mu(t) \cdot k(t)] = 0 \tag{A.75}$$

Equation (A.74) and the transition relation in part (a) of equation (A.71) form a system of ODEs in which $\dot{\mu}$ and $\dot{k}$ depend on μ, k, and c. Equation (A.73) relates μ to c, so that we can eliminate one of these two variables from the system. If we eliminate μ and take logs and time derivatives of equation (A.73), we get

$$-\rho - \dot{c}/c = \dot{\mu}/\mu$$

We can substitute this result into equation (A.74) to eliminate $\dot{\mu}/\mu$ to get

$$\dot{c}/c = (\alpha k^{\alpha-1} - \rho - \delta) \tag{A.76}$$

This condition says that consumption accumulates at a rate equal to the difference between the net marginal product of capital, $\alpha k^{\alpha-1} - \delta$, and the discount rate, ρ.

Part (a) of equation (A.71) and equation (A.76) form a system of nonlinear ODEs in k and c. In the steady state, the term $\alpha k^{\alpha-1}$ equals $\rho + \delta$, which determines the steady-state capital stock as $k^* = [(\rho + \delta)/\alpha]^{-1/(1-\alpha)}$. Part ($a$) of equation (A.71) then determines the steady-state level of consumption as $c^* = (k^*)^\alpha - \delta k^*$. Equation (A.74) implies that, as t goes to infinity, $\dot{\mu}/\mu$ tends to $-\rho$, so that $\mu(t)$ tends to $\mu(0) \cdot e^{-\rho t}$. The transversality condition in equation (A.75) can therefore be expressed as

$$\lim_{t\to\infty} [e^{-\rho t} \cdot k(t)] = 0 \tag{A.77}$$

Equation (A.77) provides a terminal condition, which, together with the initial condition $k(0) = 1$, yields the exact solution to the system of ODEs.

If we set $\rho = 0.06$, $\delta = 0$, and $\alpha = 0.3$, this system corresponds to the nonlinear system that we studied in section A.2.3 with equations (A.22) and (A.23) and linearized later in that section with equation (A.38). We know from before that this system exhibits saddle-path stability, and the initial and terminal conditions ensure that the economy starts exactly on the stable arm. We use more complicated versions of this model in the text.

Finally, we can verify that the preceding conditions imply that the steady-state value of the Hamiltonian is 0, as implied by equation (A.70):

$$\begin{aligned}\lim_{t\to\infty} [H(t)] &= \lim_{t\to\infty} \{e^{-\rho t} \cdot \log[c(t)]\} + \lim_{t\to\infty} \{\mu(t) \cdot [k(t)^\alpha - c(t) - \delta \cdot k(t)]\} \\ &= \log(c^*) \cdot \lim_{t\to\infty} (e^{-\rho t}) + 0 \cdot \lim_{t\to\infty} [\mu(t)] = 0 + 0 = 0\end{aligned}$$

Hence, although equation (A.70) is a necessary condition for optimization, it is already implied by the other conditions.

A.3.9 Transversality Conditions in Infinite-Horizon Problems

The transversality condition in equation (A.75) is not universally accepted as a necessary condition for the infinite-horizon problem. Halkin (1974) provides an example in which the optimum does not satisfy the transversality condition.[18] An even more famous counterexample is the neoclassical growth model of Ramsey (1928). The difference between the original Ramsey model and the one described in the last section is that Ramsey assumed no discounting. His version of the model is

$$
\begin{aligned}
&U(0) = \int_0^\infty \log[c(t)] \cdot dt \\
&(a)\ \ \dot{k}(t) = [k(t)^\alpha - c(t) - \delta \cdot k(t)] \\
&(b)\ \ k(0) = 1 \\
&(c)\ \ \lim_{t\to\infty} [k(t)] \geq 0
\end{aligned}
\tag{A.78}
$$

The main difference from before, equation (A.71), is that ρ has now been set to 0. An immediate problem with equation (A.78) is that, if $c(t)$ asymptotically approaches a constant (as in the previous problem), then utility is not bounded. To solve this problem, Ramsey rewrote the integrand as the deviation from a "bliss point." This revised specification will result in bounded utility if the deviation from the bliss point approaches 0 at a fast enough rate.

We found in the previous section that steady-state consumption converged to a constant, given by $c^* = (k^*)^\alpha - \delta k^*$, where k^* satisfied $\alpha \cdot (k^*)^{\alpha-1} = (\rho + \delta)$. We therefore begin with the conjecture that steady-state consumption in the present model will be $\tilde{c} = \tilde{k}^\alpha - \delta\tilde{k}$, where $\tilde{k}$ satisfies $\alpha\tilde{k}^{\alpha-1} = \delta$. The corresponding Ramseylike objective function is

$$U(0) = \int_0^\infty (\log[c(t)] - \log[\tilde{c}]) \cdot dt \tag{A.79}$$

To solve the problem of maximizing $U(0)$, as given in equation (A.79), set up the Hamiltonian,

$$H(c, k, \mu) = [\log(c) - \log(\tilde{c})] + \mu \cdot (k^\alpha - c - \delta k) \tag{A.80}$$

The first-order conditions are

$$H_c = 1/c - \mu = 0 \tag{A.81}$$

$$H_k = \mu \cdot (\alpha k^{\alpha-1} - \delta) = -\dot{\mu} \tag{A.82}$$

which correspond to equations (A.73) and (A.74).

18. This example was first presented in Arrow and Kurz (1970, p. 46). They mention, however, that the idea came from Halkin, who published the result later in *Econometrica*.

If c tends to $\tilde{c}$ as t approaches infinity, equation (A.81) implies

$$\lim_{t\to\infty}[\mu(t)] = 1/\tilde{c} > 0 \tag{A.83}$$

Since $\lim_{t\to\infty}[k(t)] = \tilde{k} > 0$, it follows that $\lim_{t\to\infty}[\mu(t) \cdot k(t)] \neq 0$; hence, the usual transversality condition in equation (A.75) is violated.

The literature has a number of examples of this sort in which the standard transversality condition is not a necessary condition for optimization. Pitchford (1977) observes that all known cases involve no time discounting. Weitzman (1973) shows that, for discrete-time problems, a transversality condition analogous to equation (A.75) is necessary when there is time discounting and the objective function converges. Benveniste and Scheinkman (1982) show that this result holds also in continuous time.

All the models discussed in this book feature time discounting and an objective function that converges. We therefore assume that the transversality condition in equation (A.75) is a necessary condition for optimization in our infinite-horizon problems.

A.3.10 Summary of the Procedure to Find the First-Order Conditions

Instead of going through the whole derivation every time we encounter a dynamic problem, we shall use the following cookbook procedure.

Step one: Construct a Hamiltonian function by adding to the felicity function, $v(\bullet)$, a Lagrange multiplier times the right-hand side of the transition equation:

$$H = v(k, c, t) + \mu(t) \cdot g(k, c, t) \tag{A.84}$$

Step two: Take the derivative of the Hamiltonian with respect to the control variable and set it to 0:

$$\partial H/\partial c = \partial v/\partial c + \mu \cdot \partial g/\partial c = 0 \tag{A.85}$$

Step three: Take the derivative of the Hamiltonian with respect to the state variable (the variable that appears with an overdot in the transition equation) and set it to equal the negative of the derivative of the multiplier with respect to time:

$$\partial H/\partial k \equiv \partial v/\partial k + \mu \cdot \partial g/\partial k = -\dot{\mu} \tag{A.86}$$

Step four (transversality condition):

Case 1: Finite horizons. Set the product of the shadow price and the capital stock at the end of the planning horizon to 0:

$$\mu(T) \cdot k(T) = 0 \tag{A.87}$$

Case 2: Infinite horizons with discounting. The transversality condition is

$$\lim_{t\to\infty}[\mu(t)\cdot k(t)] = 0 \tag{A.88}$$

Case 3: Infinite horizons without discounting. The Ramsey counterexample shows that equation (A.88) need not apply. In this case, we use Michel's condition,

$$\lim_{t\to\infty}[H(t)] = 0 \tag{A.89}$$

If we combine equations (A.85) and (A.86) with the transition equation from part (*a*) of equation (A.55), we can form a system of two differential equations in the variables μ and k. Alternatively, we can use equation (A.85) to transform the ODE for $\dot{\mu}$ into an ODE for $\dot{c}$. For the system to be determinate, we need two boundary conditions. One initial condition is given by the starting value of the state variable, $k(0)$. One terminal condition is given by the transversality condition, equation (A.87), (A.88), or (A.89), depending on the nature of the problem.

A.3.11 Present-Value and Current-Value Hamiltonians

Most of the models that we deal with in this book have an objective function of the form,

$$\int_0^T v[k(t), c(t), t]\cdot dt = \int_0^T e^{-\rho t}\cdot u[k(t), c(t)]\cdot dt \tag{A.90}$$

where ρ is a constant discount rate, and $e^{-\rho t}$ is a discount factor. Once the discount factor is taken into account, the instantaneous felicity function does not depend directly on time. If the constraints are the ones assumed before, we can solve the problem by constructing the Hamiltonian,

$$H = e^{-\rho t}\cdot u(k, c) + \mu\cdot g(k, c, t)$$

In this formulation, the shadow price $\mu(t)$ represents the value of the capital stock at time t in units of time-zero utils.

It is sometimes convenient to restructure the problem in terms of current-value prices; that is, prices of the capital stock at time t in units of time-t utils. To accomplish this restructuring, rewrite the Hamiltonian as

$$H = e^{-\rho t}\cdot [u(k, c) + q(t)\cdot g(k, c, t)]$$

where $q(t) \equiv \mu(t)\cdot e^{\rho t}$. The variable $q(t)$ is the *current-value shadow price*. Define $\hat{H} \equiv He^{\rho t}$ to be the *current-value Hamiltonian:*

$$\hat{H} \equiv u(k, c) + q(t)\cdot g(k, c, t) \tag{A.91}$$

The first-order conditions are still $H_c = 0$ and $H_k = -\dot{\mu}$. They can be expressed, however, in terms of the current-value Hamiltonian and current-value prices as

$$\hat{H}_c = 0 \tag{A.92}$$

$$\hat{H}_k = \rho q - \dot{q} \tag{A.93}$$

The transversality condition, $\mu(T) \cdot k(T) = 0$, can be expressed as

$$q(T) \cdot e^{-\rho T} \cdot k(T) = 0 \tag{A.94}$$

An interesting point about equation (A.93) is that it looks like an asset-pricing formula: q is the price of capital in terms of current utility, $\hat{H}_k$ is the dividend received by the agent (the marginal contribution of capital to utility), $\dot{q}$ is the capital gain (the change in the price of the asset), and ρ is the rate of return on an alternative asset (consumption). Equation (A.93) says that, at the optimum, the agent is indifferent between the two types of investments because the overall rate of return to capital, $(\hat{H}_k + \dot{q})/q$, equals the return to consumption, ρ.

A.3.12 Multiple Variables

Consider a more general dynamic problem with n control and m state variables. Choose $c_1(t), c_2(t), \ldots, c_n(t)$ to maximize

$$\int_0^T u[k_1(t), \ldots, k_m(t); c_1(t), \ldots, c_n(t); t] \cdot dt, \quad \text{subject to}$$

$$\begin{aligned}
\dot{k}_1(t) &= g^1[k_1(t), \ldots, k_m(t), c_1(t), \ldots, c_n(t), t] \\
\dot{k}_2(t) &= g^2[k_1(t), \ldots, k_m(t), c_1(t), \ldots, c_n(t), t] \\
&\;\;\cdots \\
\dot{k}_m(t) &= g^m[k_1(t), \ldots, k_m(t), c_1(t), \ldots, c_n(t), t]
\end{aligned} \tag{A.95}$$

$$k_1(0) > 0, \ldots, k_m(0) > 0, \quad \text{given}$$

$$k_1(T) \geq 0, \ldots, k_m(T) \geq 0, \quad \text{free}$$

The solution is similar to that for one control variable and one state variable, as analyzed before. The Hamiltonian is

$$H = u[k_1(t), \ldots, k_m(t); c_1(t), \ldots, c_n(t); t] + \sum_{i=1}^{m} \mu_i \cdot g^i(\bullet) \tag{A.96}$$

The first-order necessary conditions for a maximum are

$$\partial H/\partial c_i(t) = 0, \qquad i = 1, \dots, n \tag{A.97}$$

$$\partial H/\partial k_i(t) = -\dot{\mu}_i, \qquad i = 1, \dots, m \tag{A.98}$$

and the transversality conditions are

$$\mu_i(T) \cdot k_i(T) = 0, \qquad i = 1, \dots, m \tag{A.99}$$

A.4 Useful Results in Matrix Algebra: Eigenvalues, Eigenvectors, and Diagonalization of Matrices

Given an n-dimensional square matrix A, can we find the values of a scalar α and the corresponding nonzero column vectors v, such that

$$(A - \alpha I) \cdot v = 0 \tag{A.100}$$

where I is the n-dimensional identity matrix? Note that equation (A.100) forms a system of n homogeneous linear equations (that is, the constant term is 0 for all equations). If we want nontrivial solutions, so that $v \neq 0$, then the determinant of $(A - \alpha I)$ must vanish:

$$\det(A - \alpha I) = 0 \tag{A.101}$$

Equation (A.101) defines a polynomial equation of nth degree in α and is called the *characteristic equation*. Typically, there will be n solutions to this equation. These solutions are called *characteristic roots* or *eigenvalues*.

By construction and rearrangement of equation (A.101), each eigenvalue, α_i, is associated with a vector v_i (determined up to a scalar multiple) that satisfies

$$Av_i = v_i\alpha_i, \qquad i = 1, \dots, n \tag{A.102}$$

The vector v_i is called the *characteristic vector* or *eigenvector*. For every α_i, equation (A.102) determines an $n \times 1$ column vector (A is $n \times n$, v_i is $n \times 1$, and α_i is 1×1). We can arrange these column vectors into an $n \times n$ matrix V to get

$$AV = VD \tag{A.103}$$

where V is the $n \times n$ matrix of eigenvectors, and D is an $n \times n$ diagonal matrix with the eigenvalues as diagonal elements.

If $\det(V) \neq 0$, a condition that holds if the eigenvectors are linearly independent, V can be inverted and equation (A.103) can be rewritten as

$$V^{-1}AV = D \tag{A.104}$$

In other words, if we premultiply A by the inverse of V and postmultiply it by V, we get a diagonal matrix with the eigenvalues as diagonal elements. This procedure is called *diagonalization* of the matrix A. This result is useful for solving systems of differential equations.

Intuitively, when we diagonalize a matrix, we find a set of axes (a *vector basis*) for which the linear application represented by A can be expressed as a diagonal matrix. The new axes correspond to the eigenvectors. The linear application in these transformed axes is given by the diagonal matrix of eigenvalues.

We can state two useful results. First, if all the eigenvalues are different, then the matrix of eigenvectors is nonsingular; that is, $\det(V) \neq 0$. In this case, V^{-1} exists and, hence, the matrix A can be diagonalized.

A second interesting theorem states that the determinant and trace (the sum of the elements on the main diagonal) of the diagonal matrix equal, respectively, the determinant and trace of the original matrix. This result will be useful in situations in which we want to know the signs of the eigenvalues. Suppose, for example, that A is a 2×2 matrix and we want to know whether its two eigenvalues have the same sign. If the determinant of A is negative, the determinant of D will be negative. But, since D is diagonal, its determinant is just the product of the two eigenvalues. Hence, the two eigenvalues must have opposite signs.

As an example, consider the eigenvalues, eigenvectors, and diagonal matrix associated with $A = \left[\begin{smallmatrix} 0.06 & -1 \\ -0.004 & 0 \end{smallmatrix}\right]$. Start by constructing the system of equations

$$(A - \alpha I) \cdot v = \begin{bmatrix} 0.06 - \alpha & -1 \\ -0.004 & 0 - \alpha \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0 \tag{A.105}$$

To get a nontrivial solution, where $v \neq 0$, we must have

$$\begin{bmatrix} 0.06 - \alpha & -1 \\ -0.004 & 0 - \alpha \end{bmatrix} = 0$$

This equality determines the characteristic equation $\alpha^2 - 0.06 \cdot \alpha - 0.004 = 0$, which is satisfied for two values of α: $\alpha_1 = 0.1$ and $\alpha_2 = -0.04$. The diagonal matrix associated with A is therefore

$$D = \begin{bmatrix} 0.1 & 0 \\ 0 & -0.04 \end{bmatrix}$$

To find the eigenvector associated with the positive eigenvalue, $\alpha_1 = 0.1$, substitute α_1 into equation (A.105):

$$\begin{bmatrix} 0.06 - 0.1 & -1 \\ -0.004 & -0.1 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0$$

This equation imposes two conditions on the relation between v_{11} and v_{21}: $-0.04 \cdot v_{11} - v_{21} = 0$ and $-0.004 \cdot v_{11} - 0.1 \cdot v_{21} = 0$. The second condition is linearly dependent on the first and can be ignored. The resulting solution for v_{11} and v_{21} will therefore be unique only up to an arbitrary scalar multiple of each value. If we normalize v_{11} to 1, we get $v_{21} = -0.04$. The first eigenvector is therefore $\left[\begin{smallmatrix}1\\-0.004\end{smallmatrix}\right]$.

If we repeat the procedure for $\alpha_2 = -0.04$, we find a relation between v_{12} and v_{22}: $01 \cdot v_{12} - v_{22} = 0$. If we normalize v_{12} to 1, we get $v_{22} = 0.1$, and the second eigenvector is $\left[\begin{smallmatrix}1\\0.1\end{smallmatrix}\right]$. The two eigenvectors are linearly independent, and the matrix of normalized eigenvectors is

$$V = \begin{bmatrix} 1 & 1 \\ -0.04 & 0.1 \end{bmatrix}$$

We can now check that, indeed, $V^{-1}AV = D$ by calculating the inverse of V:

$$V^{-1} = \begin{bmatrix} 0.1/0.14 & -1/0.14 \\ 0.04/0.14 & 1/0.14 \end{bmatrix}$$

It is then easy to verify that $V^{-1}AV$ is the diagonal matrix D shown earlier.

A.5 Useful Results in Calculus

A.5.1 Implicit-Function Theorem

Let $f(x_1, x_2)$ be a bivariate function in the real space. Assume that $f(\bullet)$ is twice continuously differentiable. Let $\phi(x_1, x_2) = 0$ be an equation that involves x_1 and x_2 only through $f(x_1, x_2)$ and that implicitly defines x_2 as a function of x_1: $x_2 = \tilde{x}_2(x_1)$. An example is $\phi(x_1, x_2) = f(x_1, x_2) - a = 0$, where a is a constant. The implicit-function theorem says that the slope of the implicit function, $\tilde{x}_2(x_1)$, is

$$\frac{d\tilde{x}_2}{dx_1} = -\frac{\partial f(x_1, x_2)/\partial x_1}{\partial f(x_1, x_2)/\partial x_2} \tag{A.106}$$

This result holds whether or not an explicit or closed-form solution exists for $\tilde{x}_2(x_1)$.

As an example, consider the function $f(x_1, x_2) = 3x_1^2 - x_2$ and the equation $\phi(x_1, x_2) = 3x_1^2 - x_2 - 1 = 0$. In this case, we can find an explicit function $\tilde{x}_2(x_1) = 3x_1^2 - 1$. If we apply the implicit-function theorem from equation (A.106), we get

$$d\tilde{x}_2/dx_1 = -(6x_1)/(-1) = 6x_1$$

In this example, we do not need the implicit-function theorem to compute $d\tilde{x}_2/dx_1$, because we can differentiate $\tilde{x}_2(x_1) = 3x_1^2 - 1$ directly to get $6x_1$. The theorem is useful, however, when no closed-form solution exists for $\tilde{x}_2(x_1)$.

As another example, consider $f(x_1, x_2) = \log(x_1) + 3 \cdot (x_1)^2 \cdot x_2 + e^{x_2}$ and the equation $\phi(x_1, x_2) = \log(x_1) + 3.(x_1)^2 \cdot x_2 + e^{x_2} - 17 = 0$, which implicitly defines x_2 as a function of x_1. An explicit function $\tilde{x}_2(x_1)$ cannot be found. We can, however, compute the derivative of this function by using the implicit-function theorem,

$$d\tilde{x}_2/dx_1 = -[(1/x_1) + 6x_1x_2]/[3 \cdot (x_1)^2 + e^{x_2}]$$

A multivariate version of the implicit-function theorem is also available. Let $f(x_1, \ldots, x_n)$ be an n-variate function in the real space. Assume that $f(\bullet)$ is twice continuously differentiable. Let $\phi(x_1, \ldots, x_n) = 0$ be an equation that involves $x_1, \ldots, x_n$ only through $f(x_1, \ldots, x_n)$ and that implicitly defines x_n as a function of $x_1, x_2, \ldots, x_{n-1}$: $x_n = \tilde{x}_n(x_1, \ldots, x_{n-1})$. The implicit-function theorem gives the derivatives of the implicit function $\tilde{x}_n(x_1, \ldots, x_{n-1})$ as

$$\frac{\partial \tilde{x}_n}{\partial x_i} = -\frac{\partial f(\bullet)/\partial x_i}{\partial f(\bullet)/\partial x_n}, \quad i = 1, \ldots, n-1 \tag{A.107}$$

A.5.2 Taylor's Theorem

Let $f(x)$ be a univariate function in the real space. Taylor's theorem says that we can approximate this function around the point x^* with a polynomial of degree n as follows:

$$\begin{aligned} f(x) = {} & f(x^*) + (df/dx)|_{x^*} \cdot (x - x^*) + (d^2 f/dx^2)|_{x^*} \cdot (x - x^*)^2 \cdot (1/2!) \\ & + (d^3 f/dx^3)|_{x^*} \cdot (x - x^*)^3 \cdot (1/3!) + \cdots \\ & + (d^n f/dx^n)|_{x^*} \cdot (x - x^*)^n \cdot (1/n!) + R_n \end{aligned} \tag{A.108}$$

where $(d^n f/dx^n)|_{x^*}$ is the nth derivative of f with respect to x evaluated at the point x^*, $n!$ is n factorial $[n \cdot (n-1) \cdot \cdots \cdot 2 \cdot 1]$, and R_n is a residual. The expression in equation (A.108), with R_n omitted, is the *Taylor-Series expansion* of $f(x)$ around x^*. The presence of the residual R_n in the equation indicates that the Taylor expansion is not an exact formula for $f(x)$. The content of the theorem is that it describes conditions under which the approximation gets better as n increases.

We can check on the accuracy of the Taylor formula—that is, the size of R_n—by computing the approximation to a polynomial. If the formula is useful, it should reproduce the exact polynomial. For example, if we use a polynomial of degree 3 to approximate x^3

around 1, we get

$$\begin{aligned} x^3 &= 1^3 + (3 \cdot 1^2) \cdot (x-1) + (6 \cdot 1) \cdot (x-1)^2/2 + 6 \cdot (x-1)^3/6 + R_3 \\ &= 1 + (3x - 3) + 3 \cdot (x^2 - 2x + 1) + (x^3 - 3x^2 + 3x - 1) + R_3 \\ &= x^3 \end{aligned}$$

The residual, R_3, is 0 in this case.

As another example, we can use a polynomial of order 4 to approximate the nonlinear function e^x around 0:

$$\begin{aligned} e^x &= e^0 + e^0 \cdot x + e^0 \cdot (x^2/2) + e^0 \cdot (x^3/6) + e^0 \cdot (x^4/24) + R_4 \\ &= 1 + x + x^2/2 + x^3/6 + x^4/24 + R_4 \end{aligned}$$

The approximation (the formula with R_n omitted) gets better as the value of n increases.

If we use a polynomial of order 1 to approximate a function around a point x^*, we say that we *linearize* the function around x^*. We can also *log-linearize* a function $f(x)$ by using a first-order Taylor expansion of $\log(x)$ around $\log(x^*)$. Log-linearizations are used frequently in this book and are often useful for empirical analyses.

The two-dimensional version of Taylor's theorem is as follows. Let $f(x_1, x_2)$ be a twice continuously differentiable real function. We can approximate $f(x_1, x_2)$ around the point (x_1^*, x_2^*) with a second-order expansion as follows:

$$\begin{aligned} f(x_1, x_2) = {} & f(x_1^*, x_2^*) + f_{x_1}(\bullet) \cdot (x_1 - x_1^*) + f_{x_2}(\bullet) \cdot (x_2 - x_2^*) \\ & + (1/2) \cdot [f_{x_1 x_2}(\bullet) \cdot (x_1 - x_1^*)^2 + 2 \cdot f_{x_1 x_2}(\bullet) \cdot (x_1 - x_1^*) \cdot (x_2 - x_2^*) \\ & + f_{x_2 x_2}(\bullet) \cdot (x_2 - x_2^*)^2] + R_2 \end{aligned} \tag{A.109}$$

where $f_{x_i}(\bullet)$ is the partial derivative of $f(\bullet)$ with respect to x_i evaluated at (x_1^*, x_2^*), and $f_{x_i x_j}(\bullet)$ is the second partial derivative of $f(\bullet)$ with respect to x_i and x_j evaluated at (x_1^*, x_2^*). The linear approximation of $f(\bullet)$ around (x_1^*, x_2^*) is given by the first three terms of the right-hand side of equation (A.109).

A.5.3 L'Hôpital's Rule

Let $f(x)$ and $g(x)$ be two real functions twice continuously differentiable. Suppose that the limits of both functions as x approaches x^* are 0; that is,

$$\lim_{x \to x^*} [f(x)] = \lim_{x \to x^*} [g(x)] = 0$$

Imagine that we are interested in the limit of the ratio, $f(x)/g(x)$, as x approaches x^*. In this case, the ratio takes on the indeterminate form $0/0$ as x tends to x^*. L'Hôpital's rule is

$$\lim_{x \to x^*} \left(\frac{f(x)}{g(x)} \right) = \lim_{x \to x^*} \left(\frac{f'(x)}{g'(x)} \right) \tag{A.110}$$

provided that the limit on the right-hand side exists. If the right-hand side still equals $0/0$, we can apply l'Hôpital's rule again, until we get a result that is hopefully not an indeterminate form. L'Hôpital's rule applies to the indeterminate form $0/0$ and also works for the indeterminate form ∞/∞. The rule does not apply, however, if $f(x)/g(x)$ tends to infinity as x approaches x^*.

As an example, consider $f(x) = 2x$ and $g(x) = x$. The limit of the ratio $f(x)/g(x)$ as x tends to 0 is

$$\lim_{x \to x^*} \left(\frac{f(x)}{g(x)} \right) = \frac{0}{0} = \lim_{x \to x^*} \left(\frac{f'(x)}{g'(x)} \right) = \frac{2}{1} = 2$$

A.5.4 Integration by Parts

To integrate a function by parts, note that the formula for the derivative of a product of two functions of time, $v_1(t)$ and $v_2(t)$, implies

$$d[v_1 v_2] = v_2 \cdot dv_1 + v_1 \cdot dv_2$$

where $dv_1 = v_1'(t) \cdot dt$ and $dv_2 = v_2'(t) \cdot dt$. Take the integral of both sides of the above equation to get

$$v_1 v_2 = \int v_2 \cdot dv_1 + \int v_1 \cdot dv_2$$

Rearrange to get the formula for integration by parts:

$$\int v_2 \cdot dv_1 = v_1 v_2 - \int v_1 \cdot dv_2 \tag{A.111}$$

As an example, compute the integral $\int te^t \, dt$. Define $v_1 = t$ and $dv_2 = e^t \, dv$. By integrating dv_2 we get $v_2 = e^t$. Take the derivative of v_1 to get $dv_1 = 1$. Use the formula for integration by parts in equation (A.111) to get

$$\int te^t dt = te^t - \int e^t dt = e^t \cdot (t - 1)$$

A.5.5 Fundamental Theorem of Calculus

Let $f(t)$ be continuous in $a \leq t \leq b$. If $F(t) = \int f(t) \cdot dt$ is the indefinite integral of $f(t)$, so that $F'(t) = f(t)$, the definite integral is

$$\int_a^b f(t)\,dt = \int_a^b F'(t)\,dt = F(b) - F(a) \tag{A.112}$$

An interpretation of a definite integral is that it represents the area below the function $f(t)$ and between the points a and b (see figure A.14).

A.5.6 Rules of Differentiation of Integrals

Differentiation with Respect to the Variable of Integration The condition $F'(t) = f(t)$ implies that the derivative of an indefinite integral with respect to the variable of integration, t, is the function $f(t)$ itself:

$$\frac{\partial}{\partial t}\left(\int f(t)\,dt\right) = \frac{\partial}{\partial t}[F(t)] = F'(t) = f(t) \tag{A.113}$$

Leibniz's Rule for Differentiation of Definite Integrals Let $F(a, b, c)$ be the function describing the definite integral of $f(c, t)$, where a and b are, respectively, the lower and upper limits of integration, and c is a parameter of the function $f(\bullet)$:

$$F(a, b, c) = \int_a^b f(c, t) \cdot dt \tag{A.114}$$

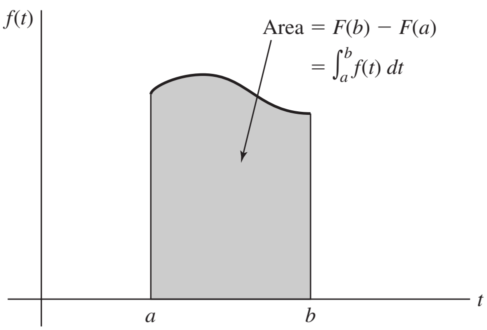


Figure A.14
The definite integral. The definite integral corresponds to the area under a curve between the limits of integration.

We assume that $f(c,t)$ has a continuous partial derivative with respect to c, $f_c(\bullet) \equiv \partial f(\bullet)/\partial c$. The derivative of $F(\bullet)$ with respect to c is

$$\frac{\partial F(\bullet)}{\partial c} = \int_a^b f_c(c,t)\,dt \tag{A.115}$$

The derivatives of $F(\bullet)$ with respect to the limits of integration are

$$\frac{\partial F(\bullet)}{\partial b} = \frac{\partial}{\partial b}\left\{\int_a^b f_c(c,t)\,dt\right\} = f(c,t)\mid_{t=b} = f(c,b) \tag{A.116}$$

$$\frac{\partial F(\bullet)}{\partial a} = \frac{\partial}{\partial a}\left\{\int_a^b f_c(c,t)\,dt\right\} = -f(c,t)\mid_{t=a} = -f(c,a) \tag{A.117}$$

We can combine equations (A.115)–(A.117) to get Leibniz's rule of integration. Suppose that a and b are functions of c:

$$F(c) = \int_{a(c)}^{b(c)} f(c,t)\cdot dt \tag{A.118}$$

Leibniz's rule is

$$\frac{dF(c)}{dc} = \int_{a(c)}^{b(c)} f_c(c,t)\cdot dt + f(c,b[c])\cdot b'(c) - f(c,a[c])\cdot a'(c) \tag{A.119}$$